4-23-2019

Understanding Multilevel Protein Communication of the Mitochondrial Carrier Family and the TIM23 Complex

Sally Chamberland
University of Connecticut - Storrs, sally.chamberland@uconn.edu

Understanding Multilevel Protein Communication of the Mitochondrial Carrier Family and the TIM23 Complex

Sally Rose Chamberland, PhD
University of Connecticut, 2019

Mitochondria are dynamic organelles that require an intricate network of protein communication that traverse multiple compartments and membrane bilayers but membrane proteins are notoriously difficult to study with traditional structure determination methods. This thesis investigates intra- and inter- membrane protein communication including residue-residue communications within a protein, between protein domains, as well as inter-protein communication between proteins of a complex and of the electron transport chain (ETC).

The first part of this thesis utilizes coevolution analysis to determine novel insights into the structure and structural dynamics of the mitochondrial carrier family (MCF) and Tim23. The ADP/ATP Carrier (AAC) is the most abundant and widely studied transporter of the MCF and imports ADP into the mitochondrial matrix and exports ATP to the cytosol by an alternating access mechanism resulting in two alternate conformational states: one with the channel interior exposed to the cytosol and one exposed to the matrix. Coevolution analysis of the AAC identifies novel residue interactions integral to the transition and stabilization of the AAC.

Tim23 is the core channel-forming protein of the translocase of the inner membrane (TIM23) complex that is responsible for the translocation and integration of approximately 70% of mitochondrial proteins. Tim23 is composed of a soluble N-

terminal domain in the intermembrane space and a membrane bound channel domain. Although key structural regions for complex assembly and protein translocation have been identified there is no high-resolution structure of Tim23. The coevolution analysis presented here predicts the first model of the Tim23 channel domain and predicts residue interactions that suggest novel structural dynamics.

The second part of this thesis investigates the inter-protein communication between subunits of the TIM23 complex. Assays of native protein assemblies, demonstrate how the lipid cardiolipin, supports proper TIM23 subunit association.

The final part of this thesis focuses on understanding the effect of NAPQI, an acetaminophen metabolite, on the electron transport activity of complex II. A reductionist approach, in which complex II was reconstituted into model membrane systems, was necessary to determine the concentration dependent effect of NAPQI on complex II activity.

Understanding Multilevel Protein Communication of the Mitochondrial Carrier Family and the TIM23 Complex

Sally Rose Chamberland

B.S., Union College, **2013**

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2019

APPROVAL PAGE

Doctor of Philosophy Dissertation

Understanding Multilevel Protein Communication of the Mitochondrial Carrier Family and the TIM23 Complex

Presented by

Sally Rose Chamberland, B.S.

Major Advisor

_____

Nathan N. Alder

Associate Advisor

_____

Joerg Graf

Associate Advisor

_____

Eric R. May

Associate Advisor

_____

Victoria L. Robinson

University of Connecticut
2019

Acknowledgements

I am truly so thankful to have made it this far in my academic career. I would not have been able to do so without the help and support of so many.

I would first like to thank my advisor Dr. Nathan Alder for supporting me and allowing me to work on such challenging and exciting research projects; none of this would have been possible without you. Thank you to my committee members Dr. Joerg Graf, Dr. Eric May, and Dr. Victoria Robinson, and Dr. Simon White. I have so enjoyed learning from you throughout this process. Thank you to my collaborator Dr. Sergey Ovchinnikov for all of your help with the coevolution analysis projects.

I would also like to thank past and present members of the Alder Lab who have generously helped me along the way: Adrian, Coscia, Matthew Greenwood, Dylan Laprise, Dr. Ketan Malhotra, Margaret Mayer, Dr. Arnab Modak, Emily Ng, Dr. Murugappan Sathappa, and Melissa Skoryk. I would like to specially thank Tess Decater and Dr. Christine Schwall-Pecci. Tess, thank you for the numerous conversations and brainstorming and for all your help with the coevolution analysis projects. Christine, thank you for introducing me to the lab and becoming a truly great friend. The guidance you have given me in lab and in life has meant the world to me.

Thank you to Dr. Linda Guilani for giving me the opportunity to work and collaborate on the UConn Provost Office Mini Grant: Streamlining the Introduction to Biochemistry (MCB 2000) Laboratory Experience to Improve Critical and Analytical Thinking Skills. This opportunity was truly one of my favorite experiences at UConn and

helped me develop as an educator in countless ways. You have been instrumental in shaping my graduate career.

Thank you to Dr. Stuart Duncan for the numerous professional development opportunities you have given me, for being a mentor to me throughout this process, and for helping me determine my future career path. Thank you to the faculty of the UConn Center for Excellence in Teaching and Learning for giving me opportunities to strengthen both my teaching and leadership abilities. Thank you especially to Dr. Peter Diplock for your continued support and encouragement during my time at UConn. Your advice and the opportunities you have provided me with have been invaluable.

Thank you to Dr. Michael Lynes for nominating me for the Center for Excellence in Teaching and Learning Outstanding Graduate Teaching Award. Without your support I would not have had the opportunity to be chosen for this award and it means so much to me.

Lastly, I would like to thank my friends and family. Thank you to Kate and Emily for girl's nights that have kept me sane through this process. Thank you to Anne and Melissa for second floor friendship that made everyday much better. Thank you to Katelyn, Céillie, Irina, and Jamie for your support and friendship and for being my truest soul mates. Thank you to Steph and Tony for your love, support, and for always being there when I needed a break. Thank you to Ryan, for always asking how the yeast were. Thank you to Evan and Ben, my favorite boys, for always reminding me how truly incredible and fun science is. Thank you to my brother David for being there whenever I needed to vent and always providing me with fun distractions.

Words cannot express how truly grateful I am for my fiancé, Bobby. You have been with me, supporting me, loving me, and encouraging me throughout this process and I do not know where I would be without.

Special thanks must also go to my parents, Ed and Beth. I would not have been able to succeed without your unwavering support. The example of hard work you have set for me has no doubtable pushed me throughout these years. I will continue to strive to make you proud.

Lastly, I must acknowledge my cousin Mary Beth, and my grandmother Ma, who are not here to celebrate this accomplishment with me but who I know would be so very proud of me. Mary Beth, your passion for life, no nonsense attitude, and strength are qualities that I will always value and aim to emulate. Ma, I wish you were here to ask me what the term “defending your dissertation” was. Your kindness and unending love of family has taught me what is really important in life. I miss you both everyday, but know you will continue to watch over me.

Table of Contents

Page

List of Figures and Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Appendix A

Appendix B

Abbreviations

AAC – ADP/ATP carrier

APAP – acetaminophen

BN-PAGE – blue native polyacrylamide gel electrophoresis

CBB – coomassie brilliant blue

CCCP – carbonyl cyanide 3-chlorophenylhydrazone

CDP-DG – cytidine-5'-diphosphate-1,2,-diacylglycerol

CL – cardiolipin

ETC – electron transport chain

FA – fatty acid

GlpT – glycerol-3-phosphate transporter

GREMLIN – generative regularized models of proteins

HHblits – HMM-HMM-based lightning-fast iterative sequence search

IM – inner membrane

IMS – intermembrane space

MIA – mitochondrial import and assembly

MIM – mitochondrial import machinery

MB – methylene blue

MCF – mitochondrial carrier family

MLCL – monolysocardiolipin

MSA – multiple sequence alignment

MSP – membrane scaffolding protein

mtDNA – mitochondrial DNA

mtHSP70 – mitochondrial heat shock protein 70

NAPQI – N-acetyl-p-benzoquinone

ND – nanodisc

OM – outer membrane

OXPHOS – oxidative phosphorylation system

PA – phosphatidic acid

PAM – presequence translocase-associated motor

PC – phosphatidylcholine

PE – phosphatidylethanolamine

PG – phosphatidylglycerol

PGP – phosphatidylclycerophosphate

PI – phosphatidylinositol

PS – phosphatidylserine

Romo1 – reactive oxygen species modulator 1

RT – room temperature

SAM -- sorting and assembling machinery

SDH – succinate dehydrogenase

SQR – succinate: ubiquinone oxidoreductase

Taz1 – tafazzin

TCA -- tricarboxylic acid

TIM22 – carrier translocase of the inner membrane

TIM23 – translocase of the inner membrane

TMS – transmembrane segment

TOM – translocase of the outer membrane

ΔΨ – membrane potential

Chapter 1

Introduction

1.1 Mitochondria Ultrastructure and Dynamics

Mitochondria are essential organelles for cellular function and play a central role in numerous cellular processes such as metabolism of amino acids, lipid synthesis, iron cofactor assembly and apoptosis (1-3). The Endosymbiosis Theory describes how mitochondria evolved as a result of engulfment of an α-proteobacterium by a host cell resulting in double membrane-bound organelles (4). These biological membranes are a key component for cellular viability by delineating different cellular spaces.

The outer membrane (OM) and inner membrane (IM) separate the mitochondria from the cytosol. The OM allows the free passage of ions and small molecules through pore forming proteins. The IM however is a more tightly packed barrier with any translocation across this membrane occurring through specific membrane proteins. Across the IM exists a membrane potential that is generated by the electron transport chain to facilitate ATP synthesis and a number of other processes. The import and sorting of proteins into the different subcompartments of the mitochondrion is mediated by specific protein complexes, which are described below. These membranes delineate multiple subcompartments, the intermembrane space (IMS), located between the OM and IM), the matrix, and cristae lumen. Contact sites occur where the IM is in close proximity to the OM, and this part of the IM is specified as the inner boundary membrane. The matrix is the innermost compartment of the mitochondria and home to mitochondrial protein transcription and translation along with many other biochemical processes. Cristae, invaginations of the IM into the matrix form the cristae lumen (5), and are the location of oxidative phosphorylation and house the complexes of the electron transport chain and ATP synthase (6) (Fig 1.1).

Phospholipids are amphipathic molecules with a hydrophilic polar head group and hydrophobic acyl chain tails and are the primary lipid component of all eukaryotic biomembranes. The primary phospholipids in both the IM and OM are phosphatidylcholine (PC) and phosphatidylethanolamine (PE) with the remaining lipid composition being made of phosphatidylinositol (PI), phosphatidylserine (PS), phosphatidic acid (PA), and cardiolipin (CL). However, the two membranes are distinct in the percent of CL present, with the OM containing roughly 5% CL and the IM containing 10-20% CL, depending on the species and cell type (7). The IM is also much more protein-rich than the OM. On a gram per gram basis, the ratio of protein to lipid in the IM is ~3-4:1, whereas in the OM it is ~1:1. This high ratio means that about 21% of total mitochondrial proteins are found in the IM (8). The mitochondrial proteome is extensive, containing approximately 1,500 and 1,000 proteins in man and yeast respectively (9,10). Although mitochondria have their own genome and the biosynthetic machinery for transcription and translation, only about 1% of mitochondrial proteins are encoded in mitochondrial DNA (mtDNA) (11). The majority of proteins are nuclear encoded, translated on cytosolic ribosomes, and must be imported into the mitochondria (12,13).

1.2 Lipids

As mentioned above, the mitochondrial OM and IM differ in lipid composition. These differences ensure optimal functional and structural integrity of the proteins that reside within these membranes and ensure proper association of protein complexes (14,15). Lipids can be separated into three general categories: bulk lipids, annular lipids, and non-annular lipids. Bulk lipids do not contact membrane proteins or protein

complexes whereas annular lipids preferentially accumulate in the vicinity of membrane proteins. Non-annular, or bound lipids mediate interactions between protein interfaces and subunits of protein complexes and also allow proper integration. Annular lipids exchange rapidly with bulk lipids and are often thought of as a solvating shell of lipids around a protein whereas non-annular lipids are often tightly bound or buried within a protein or complex (Fig 1.2). Not surprisingly then, non-annular lipids often have a role in protein and complex assembly and as such, their removal can affect protein structure and function (16). For example, minimizing exposure of the yeast cytochrome bc_1 to detergent was shown to maintain bound lipid and enhance protein activity (17). Another example of the lipid dependence of protein activity comes from the mitochondrial Translocase of the Outer Membrane (TOM) complex, which shows a marked decrease in activity with reduced levels of PE (18). Numerous proteins rely on a lipid called cardiolipin (CL) mainly found in energy transducing membranes like the mitochondrial IM (19,20). CL is required for a number of IM proteins for full catalytic activity; specifically, the ADP/ATP Carrier (AAC) (21), cytochrome *c* oxidase (22), and is essential for the assembly of respiratory complexes into supercomplexes and as a result is necessary for optimal activity (23-25).

1.2.1 CL Structure and effect on Membrane Properties

CL is an anionic phospholipid consisting of two phosphatidyl moieties connected by a central glycerol molecule. As a result of this modular structure, CL has two head groups and four acyl chains (Fig 1.3, left) (26). Due to its dimeric structure, the head group of CL has a much smaller cross sectional area than the acyl chains, giving CL an overall conical shape. The conical shape of CL results in increased lateral pressure that

can impart negative curvature in the membrane bilayer (27), which is critical for the structural integrity of highly curved membranes (28) such as mitochondrial cristae (29). Recently, it was determined the CL headgroup carries a negative two charge at physiological pH (26,30) and therefore results in a negative surface charge density of CL containing membranes. The negative surface charge density is instrumental is the electrostatic attraction and binding of peripheral and integral membrane proteins (31-33).

1.2.2 CL Biosynthesis Pathway

The biosynthesis pathway of CL (Fig 1.4A) can be broken up into three parts; 1) PA formation and translocation from the OM to IM, 2) biosynthesis of nascent CL from PA, and 3) CL remodeling, which consists of altering the composition of the acyl chains of CL. In yeast, the protein Ups1 facilitates the first step, by transferring PA from the OM to the IM (34). The conversion of PA to CL proceeds by four major reactions. Tam41, a peripheral membrane protein, catalyzes the first reaction in which PA is used to synthesize CDP-diacylglycerol (CDP-DG) (35-37). In the second reaction, Pgs1 forms phosphatidylglycerophosphate (PGP) by catalyzing the transfer of a phosphatidyl group from CDP-DG to glycerol-3-phosphate (38). Next, removal of the terminal phosphate from PGP by Gep4 results in the formation of phosphatidylglycerol (PG) (39). The next step of the pathway is the condensation of PG and CDP-DG to yield nascent CL, catalyzed by Crd1 (40,41). In the final part of this pathway nascent CL is remodeled resulting in mature CL. The composition of the acyl chains varies in a tissue- and organism-specific manner. Removal of one fatty acid (FA) chain by Cld1 results in the formation of monolysocardiolipin (MLCL) (Figure 1.3, right) (42,43). Next, the enzyme

tafazzin (Taz1) transfers an acyl chain from a donor phospholipid to MLCL from a donor lipid (44). This part of the cycle continues until all acyl chains of the CL molecule are of the proper composition.

1.2.3 Abnormal CL Biosynthesis and Barth Syndrome

The CL biosynthetic pathway was first elucidated in yeast, largely from a series of gene knockouts (Fig 1.4B). To begin, loss of Tam41, a protein involved in the early stages of the pathway, results in a complete loss of CL and an accumulation of PA (36). Loss of Crd1 (Δ*crd1*) results in a complete loss of CL and elevated levels of PG (45). Loss of Cld1 (Δ*cld1*) keeps nascent CL from entering the remodeling pathway, thereby altering acyl chain composition; however, this knockout does not yield a measureable phenotype in yeast (42, 43, 46). Loss of Taz1 (Δ*taz1*) in humans results in Barth syndrome. Barth syndrome and yeast Δ*taz1* mutants have decreased CL with a resulting increase in MLCL along with altered CL acyl chain composition. In humans, Barth syndrome is an x-linked recessive disease that often includes heart failure, myopathy, neutropenia, and growth retardation (47).

1.3 Protein Import and the TIM23 Pathway

Most mitochondrial proteins are directed to the mitochondria utilizing intrinsic targeting signals. The translocase of the outer membrane (TOM) functions as the main entry gate through the OM for the majority of mitochondrial targeted proteins (48). From there, proteins are directed to the correct mitochondrial compartment by various protein import pathways depending on the internal targeting signal they contain. The carrier translocase of the inner membrane (TIM22) integrates multispanning hydrophobic

carrier proteins into the IM (49,50). The sorting and assembly machinery (SAM) integrates β–barrel proteins into the OM (51). The mitochondrial import and assembly (MIA) machinery located in the IMS imports proteins containing characteristic cysteine motifs (52,53). The mitochondrial import (MIM) complex imports α-helical OM proteins (54,55). Lastly, the translocase of the inner membrane (TIM23) integrates proteins into the IM (56) and with the help of the presequence translocase-associated motor (PAM) imports proteins into the matrix (57). Proteins that are imported or integrated through the TIM23 complex contain a cleavable N-terminal presequence.

Presequences are highly variable in length, ranging between 10-100 amino acids (averaging 15-50 residues in length), and they lack evolutionary sequence conservation. However, all presequences form amphipathic α-helices with hydrophobic residues on one face and basic residues on the other (12). 70% of mitochondrial proteins contain an N-terminal presequence (58). A presequence containing protein is also known as a precursor protein. Once the precursor protein reaches the matrix the presequence is cleaved off by the mitochondrial processing peptidase resulting in the mature form of the protein (59,60).

1.3.1 Presequence Recognition by the TOM Complex

Precursor proteins travel across the OM through the TOM complex and across or into the IM through the TIM23 pathway. The majority of mitochondrial proteins, including presequence containing proteins, are first recognized by the TOM complex (48). The TOM complex consists of seven subunits. Tom40 is a β-barrel membrane protein that forms the channel through which precursor proteins traverse the OM. Tom70, Tom20 and Tom22 are the main receptor subunits whereas Tom7, Tom6, and Tom5 regulate

the assembly of the complex (61) (Fig 1.5). Tom20 and Tom22 first recognize a precursor protein when approaches the OM in the cytosol (62,63). Next, Tom40 mediates the translocation of the precursor protein into the IMS where it is recognized again by Tom22 (64,65), and then by the TIM23 complex.

1.3.2 The Subunits of the TIM23^{CORE} Complex

The TIM23 complex is versatile in that it facilitates the translocation of both matrix bound and IM integrated presequence containing proteins (56). Regardless of protein identity, translocation and integration both occur in a membrane potential dependent manner (66,67). The TIM23 complex is modular, in that it contains 11 identified subunits and assembles into multiple functional organizations depending on the final destination of the precursor protein (Fig 1.5). TIM23^{CORE}, the main functional unit, contains Tim23, Tim17, and Tim50. Tim50 has an N-terminal matrix domain, a single transmembrane segment (TMS) spanning the IM, and a large C-terminal hydrophilic domain located in the IMS. Importantly, a β-hairpin located in the IMS promotes the interaction of Tim50 with Tim23 (68). The C-terminal IMS domain of Tim50 acts as the central receptor of the complex and plays a crucial role in transferring precursor proteins from TOM to TIM23 (69-71). Deletion of Tim50 is lethal in yeast, and decreased expression levels of Tim50 results in reduced protein import.

Tim23 and Tim17 are homologous proteins that make up the core voltage-gated channel (72). Tim23 has a bipartite organization containing a soluble N-terminal IMS domain that is intrinsically disordered, and a C-terminal channel-forming domain (66, 73). The N-terminal domain is able to span the IMS and helps transfer precursor proteins from the TOM complex (74) in cooperation with Tim50. The C-terminal domain

of Tim23 is predicted to have four TMSs that constitute the ΔΨ dependent translocation channel (73). Tim17 is an integral membrane protein that lacks an extensive IMS domain like Tim23 and instead consists of just four TMSs, with the N-terminal and C-terminal domains located in the IMS (75).

1.3.3 Integration of IM Proteins

In addition to the $TIM23^{CORE}$ subunits, $TIM23^{SORT}$, an alternate modular assembly of the TIM23 complex, contains proteins Tim21 and Mgr2 (Fig 1.5). Tim21 is anchored in the IM by a single TMS, and has an N-terminal domain located in the matrix and a C-terminal domain located in the IMS. Tim21 is predicted to interact with Tom22 through electrostatic interactions (76) and works antagonistically to Tim50 by reducing the accumulation of precursor protein at the TOM complex (56). Tim21 has been shown to recruit the respiratory chain III-IV supercomplex resulting in promotion of ΔΨ driven protein translocation (77). Mgr2, a small hydrophobic protein, consisting of two TMS in the IM and a small N-terminal hydrophilic IMS domain (78), promotes the coupling of Tim21 to the $TIM23^{CORE}$ (79) and has been shown to act as a quality control checkpoint for laterally sorted IM proteins (80).

1.3.4 Import of Matrix Bound Proteins

Tim17 is thought to promote protein sorting and upon Tim21 dissociation from $TIM23^{CORE}$, Tim17 acts as a docking point for the ATP-driven PAM complex (56). The PAM complex is composed of six subunits (Fig 1.5) Tim44, mtHsp70, Pam16, Pam17, Pam18, and Mge1. When $TIM23^{CORE}$ associates with the PAM complex the $TIM23^{MOTOR}$ complex is formed. $TIM23^{MOTOR}$ is essential for the complete translocation of precursor

proteins across the IM and into the matrix (57). A presequence containing protein is driven into the matrix through interaction with the ATP hydrolyzing mitochondrial heat shock protein 70 (mtHsp70) (81). Tim44 has two domains, the N-terminal domain and C-terminal domain, and acts as a scaffold, linking $TIM23^{CORE}$ to the other PAM complex components (82). The C-terminal domain interacts with Tim23 and Tim17 (83). The N-terminal domain is intrinsically disordered and thought to play a regulatory role as it has been shown to interact directly with mtHsp70 and Pam16 (82). Pam16 is a peripheral membrane protein that contains a J-like domain with a modified HPD motif. Pam18 is a J-domain containing protein, anchored to the IM by a single pass TMS. Pam18 has a stimulatory effect on ATP hydrolysis by mtHsp70 and, as a result protein import into the matrix, which is additionally regulated by Pam16 (84-86). Mge1 is a nucleotide exchange factor that promotes the release of ADP from mtHsp70, allowing the next reaction cycle to occur (87). The final subunit of PAM is Pam17, another regulatory protein that aids in the organization of the complex, however its specific function is unknown (88-90).

Although high-resolution structures are available for some of the soluble domains of the TOM and TIM23 complex subunits (68,91,92), there is no structural information available for Tim23 or Tim17 and it is unknown how these subunits come together to form the voltage-gated pore.

1.3.5 The TIM23 Complex and CL

A number of studies have shown CL is important for the functional integrity of the TIM23 complex (35-37, 93-95). In one example, yeast cells lacking Tam41 showed decreased protein import possibly due to a decrease in association of $TIM23^{CORE}$ with

the PAM complex (37, 93). Upon analysis with BN-PAGE, yeast cells lacking Tam41 show a destabilization of the $TIM23^{SORT}$ complex. Furthermore, cells lacking Ups1 had decreased CL and decreased association of $TIM23^{CORE}$ with the PAM complex (95).

Another example is the role of CL in the membrane interactions of Tim44 (96,97). The interaction of helices A1 and A2 of the N-terminal domain of Tim44 with liposomes is enhanced when CL was present. Supporting this, deletion of helix A1 abolishes the ability of Tim44 to associate with CL containing liposomes (97) suggesting CL is important for the interaction of Tim44 with the IM.

More recently, a more complete picture of the effect of CL on the TIM23 complex was described by investigating the effect of CL on Tim50 (98). Through a number of techniques it was determined that the soluble domain of Tim50 (yeast residues 133-476) interacts with membranes in a CL-dependent way, and that this interaction then promotes the Tim50 association with the channel domain of Tim23 (98). MD simulations and biochemical experiments more specifically identified the β-hairpin containing face of Tim50 to stabilize the interaction with the membrane (98). Together, these observations suggest that CL could promote protein import through the TIM23 complex by promoting the association between Tim50 and Tim23.

The N-terminal domain of Tim23 (residues 1-97 of *S. cerevisiae*) is located in the IMS and the first 50 residues are required for yeast growth at 37°C (74). It has been shown that during protein translocation the TOM and TIM23 complexes associate transiently to form a supercomplex at contact sites (99-101). These two facts led researchers to investigate the role the IMS domain of Tim23 may play in the formation of contact sites during protein translocation. Using nuclear magnetic resonance,

residues 1-7 and 29-46 were identified to bind directly to membranes. Residues W3 – L4 – F5 have been termed the “hydrophobic hook” with W3 known to insert into the hydrophobic milieu of the OM. CL has been shown to enhance this membrane interaction (102).

1.4 The Conservation of Protein Import and Associated Diseases

1.4.1 The TOM Complex

Yeast is an effective model to study mitochondrial protein import due to the high conservation of these protein import complexes. The TOM complex is conserved from yeast to human, containing the same seven subunits (103). The only differences between yeast and human TOM arise in the cytosolic chaperones involved in delivering presequence containing proteins to the complex (104).

With the TOM complex having such a large role in protein import it is not surprising it has been linked to disease, specifically Parkinson’s disease (105). The binding of α-synuclein to Tom20 has been shown to impair the Tom20-Tom22 interactions, thereby impairing protein import (105). Additionally, PINK1, a mitochondrial kinase associated with familial cases of Parkinson’s disease, can form stable complexes with the TOM complex in damaged mitochondria. These complexes then serve as a signal of mitochondrial damage (106,107).

1.4.2 The TIM23^{CORE} Complex

The structural organization and function of TIM23^{CORE} is conserved from yeast to humans again validating the use of a yeast model for these studies (108), however there are differences. As stated above Tim50 is the first point of contact in the TIM23

complex for presequence containing proteins. Human Tim50 has been shown to have phosphatase activity (109), whereas in yeast there is no evidence of this. Furthermore, whereas yeast Tim17 is encoded by a single gene, in humans, Tim17 exists as three separate isoforms (a, b1, and b2), which are present in distinct TIM23 complexes (110).

A number of pathological conditions can be linked to defects in components of the protein import machinery. Tim23 is essential for viability in both yeast and mice (111,112). The first evidence of a pathological implication due to altered protein import was shown with human deafness dystonia syndrome. This disease is caused by mutations in Tim8, a subunit of the TIM22 pathway, which is responsible for the import of Tim23 (see section 1.6.1 for detail), indicating poor protein import and a lack of Tim23 present in the mitochondria as possible indirect causes for disease (113,114). Another example is with high expression of both Tim50 (115) and Tim17 (116,117) which has been associated with breast cancer. Additionally, the suppression of Tim50 hinders cancer cell proliferation (115).

1.4.3 The TIM23^{MOTOR} Complex

The function of TIM23^{MOTOR} is also conserved from yeast to humans, and includes the same major subunits. However, as with Tim17, humans have two orthologs of Pam18, DnaJC15 and DnaJC19. DnaJC15 associates with TIM23^{CORE} containing Tim17a whereas DnaJC19 associates with complexes containing Tim17b (110). Another difference lies within mtHsp70, the matrix-localized chaperone that serves as the central ATPase motor protein of the PAM complex. This subunit is conserved throughout all eukaryotes. However, whereas humans contain only one homolog termed

mortalin (118), yeast contain three homologs, of which the one most critical for protein import is Ssc1 (119,120).

A number of diseases have been associated with defects in the subunits of the PAM complex. A single point mutation in human Tim44 is associated with oncocytic thyroid carcinoma (121) and up regulation of Tim44 is associated with diabetes (122). Magmas, the human Pam16 ortholog, expression has been associated with prostate cancer cells (123). In addition, a mutation in Magmas has been correlated with skeletal dysplasia (124). Mortalin is mostly present in the mitochondria is also found in the cytosol and able to bind inactive p53, this results in uncontrolled cell proliferation and a number of cancers (125).

1.4.4 The TIM23^{SORT} Complex

As in yeast, human TIM23^{SORT} is composed of Tim23, Tim17, Tim50, Tim21 and Romo1 (reactive oxygen species modulator 1), a homolog of Mgr2 (126). In both humans and yeast, Tim21 interacts with both the TIM23 complex and the respiratory chain III-IV supercomplex, thereby increasing the $\Delta\Psi$ in the vicinity of the TIM23^{SORT} complex and aiding in the lateral insertion of proteins (77). In humans, Tim21 plays an additional role in the assembly of respiratory chain complexes (127). Romo1 was first discovered through the identification of up-regulated proteins in tumor tissue, which resulted in increased reactive oxygen species production (128). Overexpression of Romo1 is specifically associated with the poor prognosis of hepatocellular carcinoma patients (129).

1.4.5 Summary

Conservation of the TIM23 and TOM complexes indicates just how important protein import is to organelle survival and cell integrity. While differences remain between yeast and human, the main subunits remain conserved, making yeast a fitting model organism to study this system (130).

The number of diseases involving subunits of the TOM or TIM23 complex again emphasizes the importance of this functional system to organelle survival and cell integrity. The high number of resulting cancers and other diseases caused by improper expression of TIM23 and TOM complex subunits demonstrates just how important it is to understand the process of protein import and subunit structure to be able to combat these disease states.

1.5 The Tim23 Protein Superfamily

The Tim23 protein superfamily refers to homologous proteins Tim23, Tim17 and Tim22. Like Tim23 and Tim17, Tim22 contains four TMSs with the N- and C-terminal ends located in the IMS (49). Tim22 is not part of the TIM23 complex, but is the core subunit of the TIM22 complex.

1.5.1 The TIM22 Complex

The TIM22 complex integrates polytopic IM proteins, namely mitochondrial carrier family (MCF) proteins, like the ADP/ATP Carrier (AAC) (49,131), and the protein translocation proteins Tim23 and Tim17 (132-135). Tim22, like Tim23 and Tim17 integrates proteins into the IM in a ΔΨ-dependent manner (136). Tim54 is thought of as the receptor-like protein of the complex and recruits the chaperone complexes Tim9-

Tim10-Tim12 (137,138) and Tim8-Tim13 (139) to the TIM22 complex. These chaperone complexes help protect the hydrophobic IM proteins destined for IM integration from aggregating. Finally, the subcomplex consisting of Tim18 and Sdh3 aids in the assembly of the TIM22 complex (140).

1.5.2 Protein Integration through the TIM22 Complex

Proteins integrated into the IM through the TIM22 complex do not contain an N-terminal presequence; instead they contain multiple internal signals, which target these hydrophobic proteins to the mitochondrial IM (141,142). Proteins targeted to the TIM22 pathway also utilize the TOM complex to traverse the OM (143). Once proteins arrive in the IMS they interact with chaperone complexes that facilitate their interaction with the TIM22 complex (135,144,145). Specifically, it has been shown AAC requires Tim9-Tim10 for proper import (146) whereas both Tim9-Tim10 and Tim8-Tim13 complexes have been shown to interact with different sites of Tim23 to promote import (145).

1.6 The ADP/ATP Carrier

1.6.1 The Mitochondrial Carrier Family

The MCF is comprised of transporters that mediate the flux of small molecules between the mitochondrial matrix and the cytosol. Mitochondrial carriers are integrated into the mitochondrial IM, selectively transporting a range of substrates including nucleotides, amino acids, fatty acids, inorganic ions and cofactors. These carriers therefore play a critical role in coordinating major metabolic pathways of the mitochondria, including the citric acid cycle, oxidative phosphorylation, and fatty acid oxidation, with the rest of the cell (147). MCF transporters are α-helical bundle proteins

with six transmembrane segments. They have a common tripartite structural organization comprised of a tandem arrangement of three homologous modules, each approximately 100 amino acids in length (148). Each module consists of two transmembrane α-helices connected by shorter matrix-facing helices (Fig 1.6). Despite the sequence and structural similarities that unite MCF transporters, each has well-defined substrate specificity.

1.6.2 The AAC Structure

The AAC, the most abundant and widely studied member of the MCF, mediates the bidirectional and electrogenic transport of adenine nucleotides, importing ADP into the matrix and exporting ATP to the cytosol (149). This transporter is therefore essential for energy balance within eukaryotic cells. During its transport cycle, the AAC is believed to have two main conformational states: the c-state and the m-state, in which the substrate binding site is accessible to the cytosolic- and matrix-facing sides of the mitochondrial IM, respectively (150-154). Our current knowledge of the structures of MCF transporters is based primarily on the AAC given the availability of high-resolution crystal structures of the AAC in the c-state from bovine heart (155,156) and from the yeast, *S. cerevisiae* (157), as well as projection structures of yeast AAC in lipid bilayers (158). More recently, the crystal structure of the AAC in the m-state from *Thermothelomyces thermophilia* was solved providing even more detail into the structure of the AAC (159) and as a result other MCF transporters.

To date, all AAC structures have been solved in an inhibitor-stabilized state. The inhibitor, carboxyatractyloside, maintains the transporter in the c-state (160), whereas the inhibitor bongkrekic acid maintains the transporter in the m-state (161). These AAC

structures reveal a pseudo threefold symmetry consistent with the canonical tripartite domain organization of MCF transporters. Its transmembrane helices are highly tilted relative to the bilayer normal, forming a barrel-like structure with a deep conical substrate-binding cavity exposed to the cytosol in the c-state and the matrix in the m-state (Fig 1.6B,C). Closure of the barrel is attained through a gate that is about 15Å thick on either the cytosolic or matrix side of the protein in the m-state and c-state respectively (159).

In the c-state the closure of the barrel is attained by the odd-numbered transmembrane helices (H1, H3, and H5). These helices contain the MCF signature motif Px[D/E]xx[K/R] (Fig 1.6A, sm1), which is conserved among all members of the MCF (162). The proline residue of this motif induces a sharp kink within the odd-numbered helices, causing the C-terminal ends of these helices to converge at the base of the binding cavity. The interaction of these kinked helices is stabilized by an intermodule salt bridge network formed by the charged residues of the MCF motif (155-157) and braced by glutamate side chains (157). Situated on the matrix side of this electrostatically-stabilized network are three amphipathic helices (h12, h34, and h56) that connect the transmembrane helices of each module. Together these matrix helices are believed to impart stability to the c-state of the AAC by maintaining closure on the matrix side of the protein and the formation of the matrix gate.

In the m-state the even-numbered helices (H2, H4, H6) are important for the cytosolic gate closure and conformational stabilization. The even-numbered helices contain a conserved motif, [Y/F][D/E]xx[K/R] (Fig 1.6A, sm2) that forms a complementary intermodule salt bridge network due to the charged residues of the motif

(157,163,164) that is braced by tyrosine side chains (159). Bulky hydrophobic and aromatic residues form a hydrophobic plug located on the cavity side of the salt bridge network supports the cytosolic gate closure (159).

These salt bridge networks and residue braces result in the closure of the substrate-binding cavity on the cytosolic side or matrix side of the protein. In both the c-state and m-state there are approximately equal interaction energies and a similar number of polar interactions (159) suggesting comparable stability of both conformational states of the protein.

1.6.3 The AAC Transport Mechanism

It is generally accepted that the AAC and other MCF transporters function by an alternating access mechanism (165). In this model for solute transport, substrate binding to a site accessible from one side of the membrane leads to a conformational change in the transporter that exposes the substrate to the opposite side of the membrane, allowing the substrate to dissociate, in this way the substrate binding site is thought of as a fulcrum (159). In such mechanisms, the interconversion of these extreme conformers proceeds via one or more intermediary structural states (Fig 1.6D) that are still unresolved. Because available crystal structures of AAC are all locked in either the c-state or m-state, the structural basis by which the carrier can convert between these states by this alternating access mechanism to catalyze the equimolar exchange of ADP and ATP remains an open question. However, several lines of evidence offer insights into possible models. First, measurements of transport activity support a model in which AAC functions as a monomeric unit (166-167) with a single substrate binding site whose location is common to all MCF transporters (168) which

has been supported by photoaffinity labeling (169,170) and site-directed mutagenesis experiments (171). Second, the nature of substrate binding is well established. In the c-state, the conical binding cavity extends ~30Å to the center of the membrane and is lined by patches of conserved basic and aromatic residues that coordinate the phosphate and adenine/ribose moieties of ADP, respectively (168, 172). Molecular dynamics (MD) simulations based on the bovine AAC structure show that the cavity has a strong electrostatic potential that drives ADP inward, where ADP phosphates form salt bridges primarily with residues K22, R79, R279, and R235 (173,174). Third, the AAC is known to be structurally dynamic. Studies using hydrogen/deuterium exchange coupled with mass spectrometry suggest that AAC displays greater solvent accessibility on the cytoplasmic- and matrix-facing sides when locked in the c- and m-states, respectively (175,176) and NMR relaxation experiments show that the AAC undergoes highly asymmetric structural deviations from the c-state in an ADP-dependent manner (177).

1.6.4 The AAC and CL

Initial purification and activity assays suggested the importance of CL in AAC activity (178,179). Previous studies of AAC have identified three CL molecules that bind tightly to both bovine and yeast AAC (155-157). Observations of CL binding sites identified two conserved glycine residues that play a key role in binding (157, 180). CL interacts with the AAC through motifs found on the N-terminal ends of the even-numbered helices, [Y/W/F][R/K]G, and matrix helices, [Y/F]xG. The conserved glycine residues provide breaks in the TMSs allowing the phosphate groups of CL access to form hydrogen bonds with the positively charged dipole ends of the even-numbered and matrix helices. The bulky hydrophobic residues of these motifs are then able to interact

with the acyl chains of CL, however no specific conformation is adopted (157, 180). More recently molecular dynamics (MD) simulations showed agreement with the three observed binding sites of CL identified in the crystal structure and demonstrates the dynamic nature of these interactions (181,182). While the role of CL in the function of the AAC is still unclear, the conserved nature of the binding motifs stresses the importance of these sites.

1.7 Research Goals

In this thesis I have used a number of techniques to understand the necessity of proper inter- and intra- protein communication at multiple levels. This includes communication between protein residues and domains and protein-protein interactions both within complexes or cellular processes (Fig 1.7).

In chapters two and three, coevolution analysis was used to determine the most important evolutionary constraints on a protein and the residues that communicate within a protein to necessitate proper protein structure and function. This research demonstrated the strength of coevolution analysis to predict protein structure and elucidated previously unknown structural dynamics and evolutionary important interactions of not only the AAC but of MCF proteins in general. I then utilized this technique to understand more about the elusive structure of the channel domain of Tim23 and residues that are potentially important for the interactions and protein communication between Tim23 and Tim17.

In chapter four, I expanded on understanding inter-protein communication by investigating the assembly states of the TIM23 complex. Due to the modular nature of

the TIM23 complex it has been shown to exist in multiple assembly states when analyzed by BN-PAGE. I investigated the role of lipids, specifically CL, and the energized state of the IM in mediating these protein-protein interactions. To better understand this modular complex, I take this research one step further by determining if these effects are reversible through CL supplementation studies. This readout of TIM23 assembly states also acts as a readout to investigate how model membrane systems could be utilized as a mitochondria targeted lipid delivery system.

Finally, chapter five addresses the necessity for proper protein communication to support oxidative phosphorylation, a vital cellular process. This study aimed to understand the mechanism of compromised protein communication, specifically the inhibition of respiratory complex II. A reductionist system was necessary to understand the underlying mechanism of acetaminophen induced liver injury and the effect of an oxidative metabolite of acetaminophen on oxidative phosphorylation. My work confirmed the effect of this metabolite by using a complex II specific activity assay that allowed clear understanding of the effect. With this mechanism more completely understood my collaborators continued the study to determined how to circumvent this compromised process and restore proper electron transport chain communication.

1.8 Figures

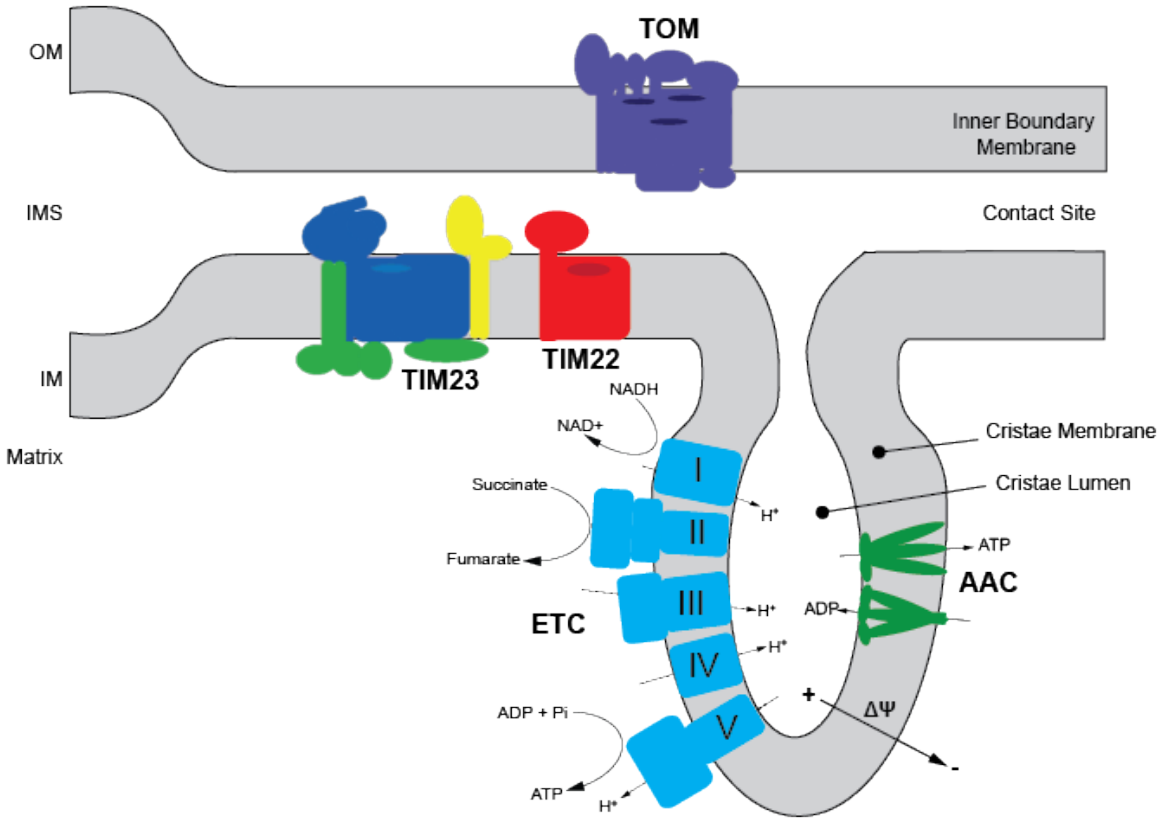


Figure 1.1. Ultrastructure of mitochondria. Mitochondria contain two membranes, the OM and IM, which delineate multiple aqueous spaces including the IMS and matrix. The TOM complex is located in the OM and the TIM23 and TIM22 complexes are located in the IM. TOM, TIM23 and TIM22 are all involved in protein transport. Invaginations of the IM into the matrix are known as cristae which are the location of the ETC and the AAC.

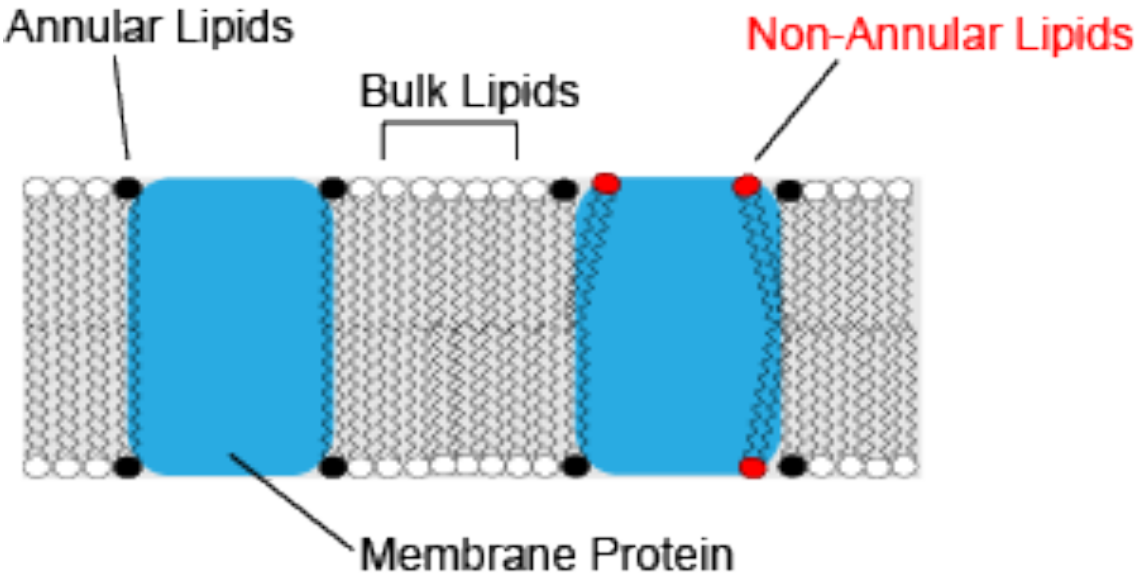


Figure 1.2. Lipid – protein association. Schematic demonstrating the different types of lipids and lipid – protein interactions. Membrane proteins shown as blue rectangles. Lipid headgroup colors indicate the type of lipid. Annular lipids shown in black. Bulk lipids shown in white. Non-annular or bound lipids shown in red.

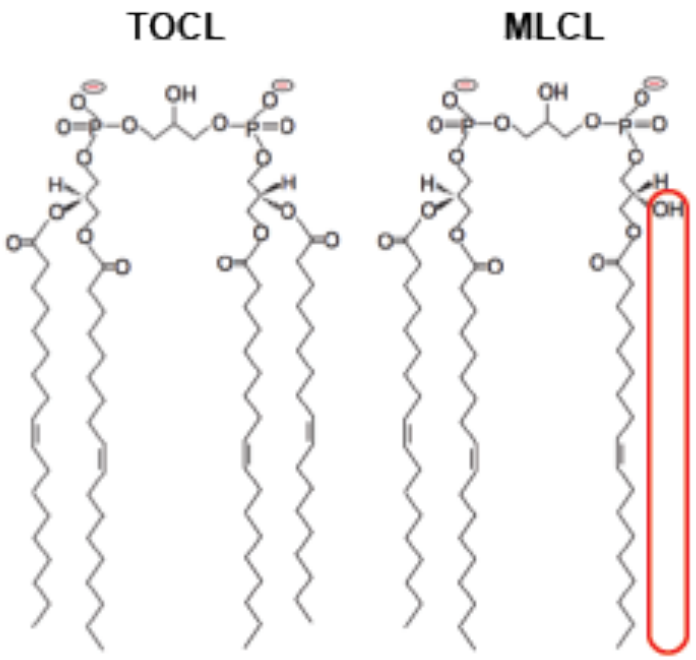


Figure 1.3. Chemical structures of tetraoleoylcardiolipin (TOCL) and monolysocardiolipin (MLCL). Structures of TOCL (*left*) and MLCL (*right*) are shown with oleoyl-based (18:1) acyl chains.

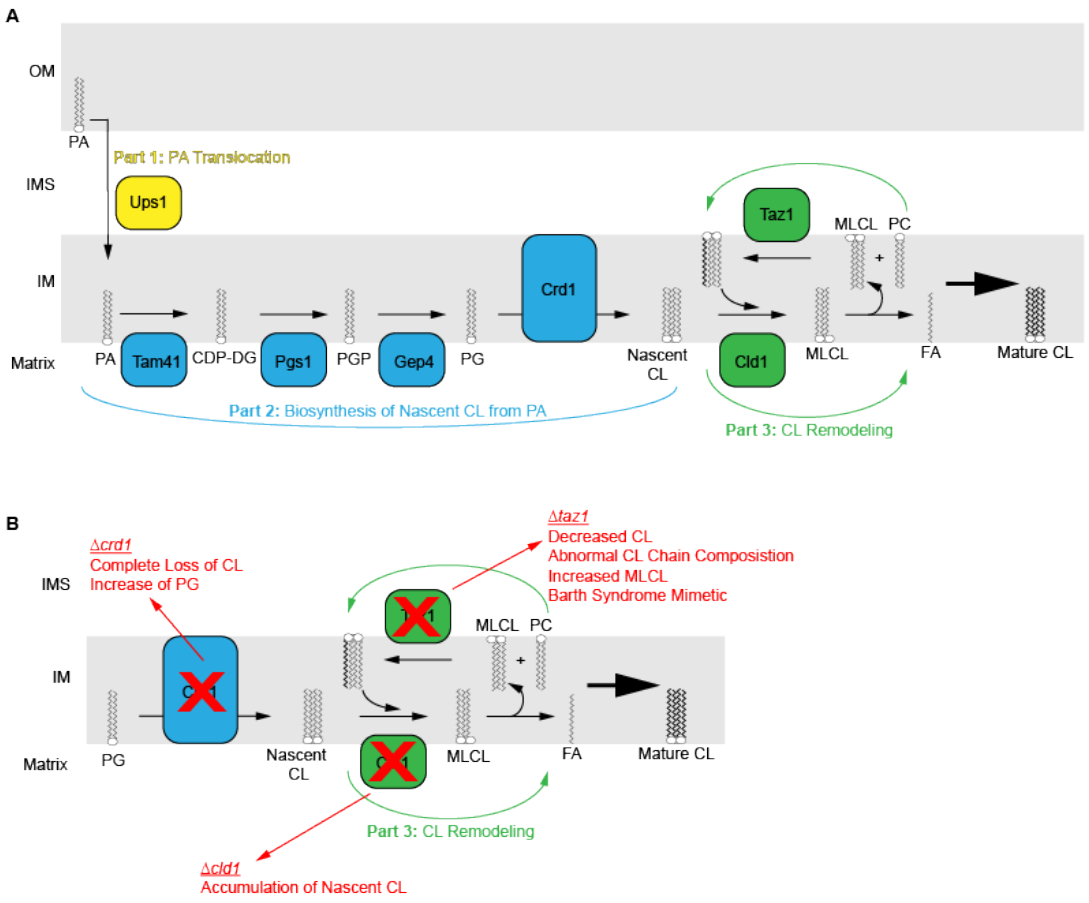


Figure 1.4. CL biosynthesis pathway. (A) CL biosynthesis occurs in three parts. Part one requires Ups1 and the translocation of PA from the mitochondrial OM to the IM. Part two utilizes multiple proteins to catalyze the conversion of PA to nascent CL. In the final part, CL remodeling, the acyl chains of nascent CL are exchanged until proper acyl chain composition is achieved and mature CL is formed. (B) Yeast strains exist to research how abnormal CL biosynthesis affects lipid composition. Specifically, the Δ*taz1* yeast strain is a model of the human disease Barth Syndrome.

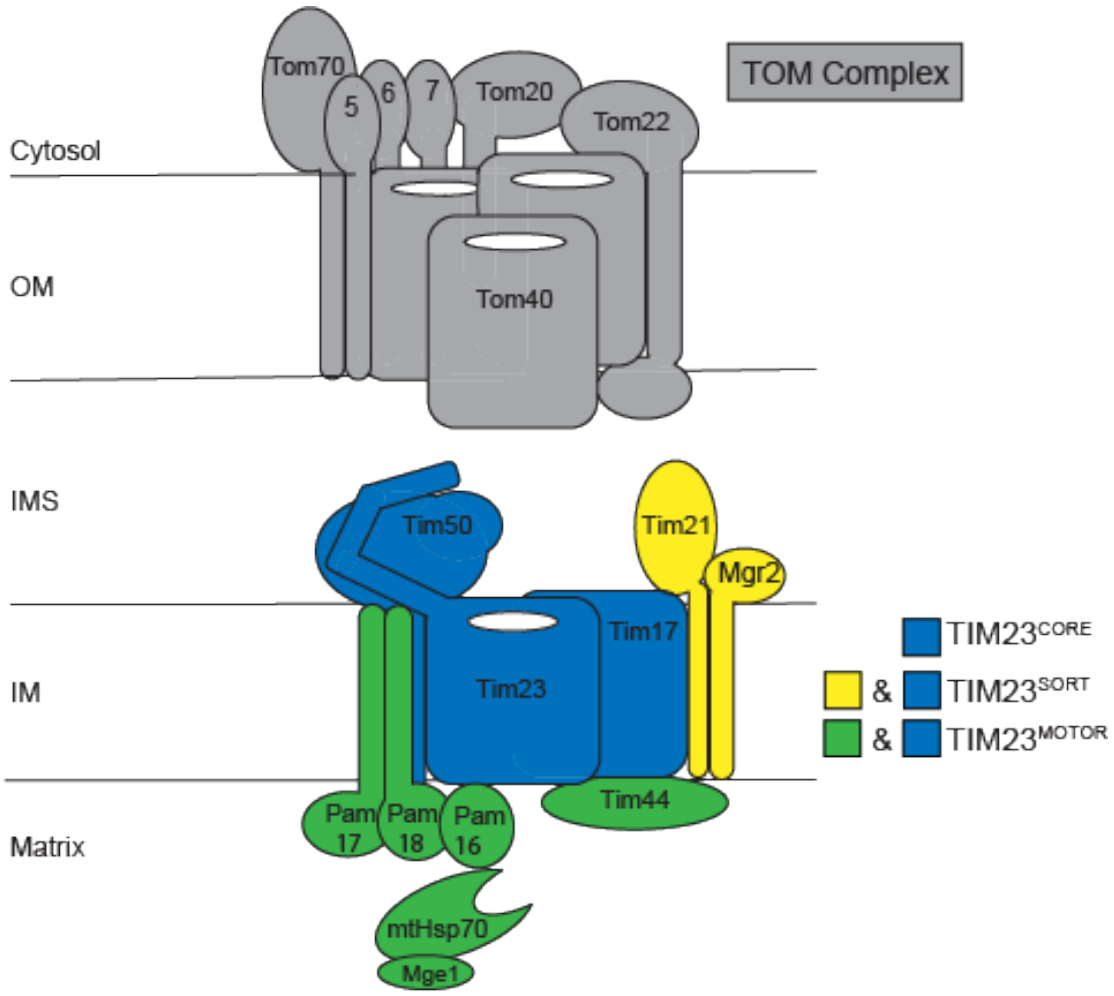
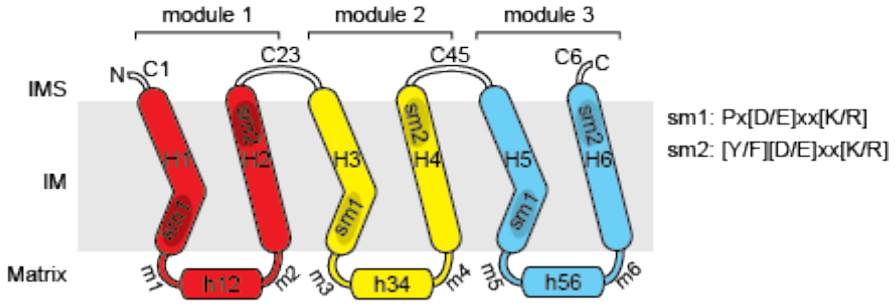
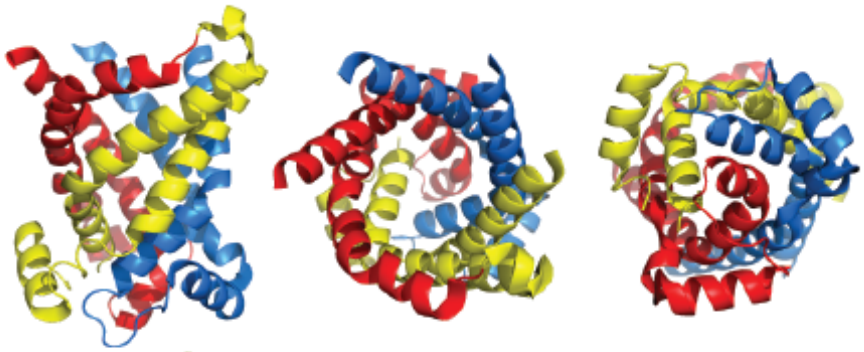


Figure 1.5. Schematic of the TOM and TIM23 complexes. The TOM complex is located in the mitochondrial OM and is composed of seven different subunits, see text for more detail. The TIM23 complex can assemble into different forms depending on the protein being transported. The $TIM23^{CORE}$ complex (blue) consists of Tim50, Tim23, and Tim17. The $TIM23^{SORT}$ complex (blue and yellow), assembles when integrating proteins into the IM, consists of $TIM23^{CORE}$ and additional subunits Tim21 and Mgr2. Finally the $TIM23^{MOTOR}$ complex (blue and green), assembles when translocating proteins into the mitochondrial matrix and consists of $TIM23^{CORE}$ and the PAM complex subunits.

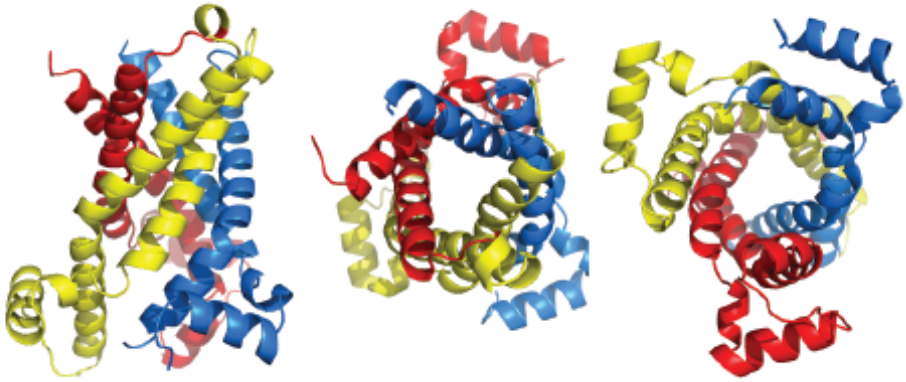
A


module 1
module 2
module 3
N
C1
C23
C45
C6
C
IMS
IM
Matrix
H1
H2
H3
H4
H5
H6
sm1
sm2
h12
h34
h56
m1
m2
m3
m4
m5
m6
sm1: Px[D/E]xx[K/R]
sm2: [Y/F][D/E]xx[K/R]

B



C



D

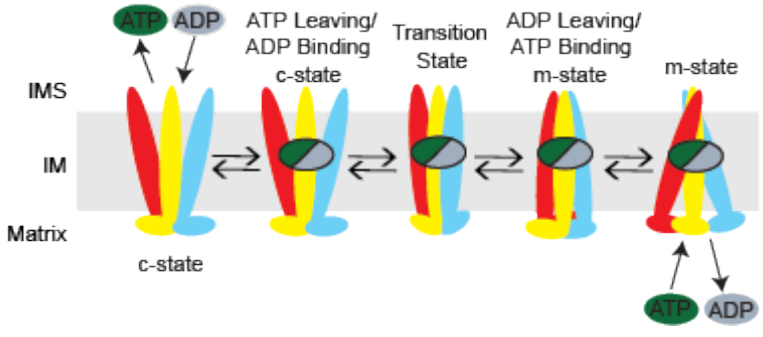

ATP
ADP
ATP Leaving/
ADP Binding
c-state
Transition
State
ADP Leaving/
ATP Binding
m-state
m-state
IMS
IM
Matrix
c-state
ATP
ADP

Figure 1.6. The structure and function of the ADP/ATP carrier. (A) Domain organization and topology of AAC. Individual modules are identified by color with general locations of conserved sequence motifs (sm1 and sm2). The nomenclature used here for the structural elements is shown, including transmembrane helices (H1-6), matrix helices (h12, h34, h56), cytosolic-facing loops (C1, C23, C45, C6), and matrix-facing loops (m1-6). (B) Crystal structure of the c-state conformation of the AAC (PDB: 4C9G). *Left*, side view. *Middle*, cytosolic view. *Right*, matrix view. Modules are color coded as in panel A. (C) Crystal structure of the m-state conformation of the AAC (PDB: 6GCI). *Left*, side view. *Middle*, cytosolic view. *Right*, matrix view. Modules are color coded as in panel A. (D) Predicted transport cycle of AAC. AAC transports either a single ADP from the cytosol into the mitochondrial matrix or a single ATP from the matrix into the cytosol by a proposed alternating mechanism with two alternative conformations: the c-state, open to the cytosol, and the m-state, open to the matrix.

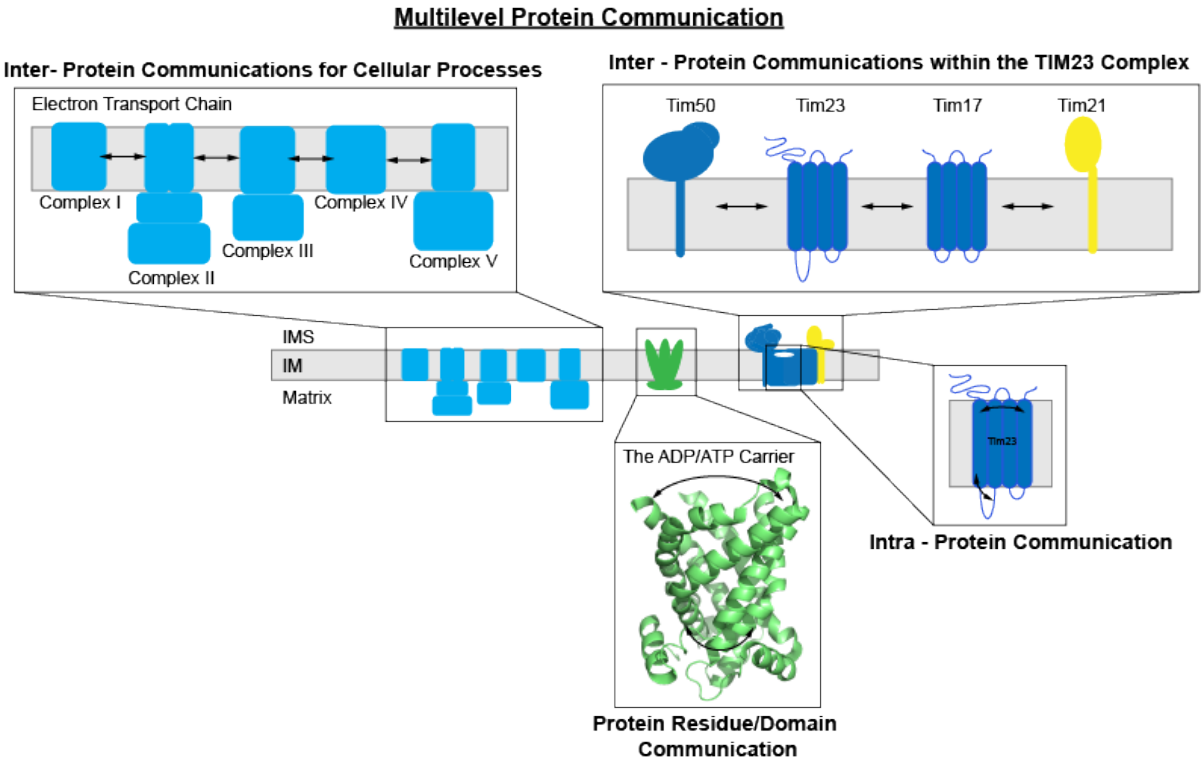


Figure 1.7. Research goals. This thesis investigates multi-level inter- and intra-protein communication throughout the mitochondria. Protein communication studied, depicted by black arrows. Protein-protein communications for cellular processes and complex formation is studied through the electron transport chain and the TIM23 complex. Intra-protein communication is studied through Tim23 and the AAC.

Chapter 2

Coevolution Analysis of the Mitochondrial Carrier Family Provides Insight into Conformational Dynamics and Transport Mechanisms

Derived from:

Chamberland SR., Decater TA., Ovchinnikov S., Alder NN. Coevolution Analysis of the Mitochondrial Carrier Family Provides Insight into Conformational Dynamics and Transport Mechanisms, *In preparation*.

2.1 Abstract

The MCF is comprised of transporters that mediate the flux of small molecules between the mitochondrial matrix and the cytosol. These carriers play a crucial role in coordinating major metabolic pathways of the mitochondria by transporting a range of substrates including nucleotides, amino acids, fatty acids, and inorganic ions and cofactors. The AAC is the most abundant and widely studied transporter of the mitochondrial carrier family. The AAC resides in the mitochondrial IM where it imports ADP into the mitochondrial matrix and exports ATP to the cytosol, thereby playing a critical role in the energy balance of eukaryotic cells. The AAC is a hallmark example of MCF structure, consisting of three homologous modules, each with two transmembrane helices and one matrix-facing loop. The MCF proteins transport substrates across the IM by an alternating access mechanism resulting in two alternate conformational states: one with the channel interior exposed to the cytosol (the c-state) and one exposed to the matrix (the m-state). Structural studies of the AAC have confirmed these structural states and high-resolution structures of the AAC locked in the c-state and m-state have been published. Key features of this structure include highly tilted transmembrane helices with sharp bends in odd-numbered helices, forming a barrel-like structure that is stabilized by a network of electrostatic interactions in both conformational states. Yet because inhibitors have been used that lock this protein in either conformational state, many critical questions about the transport mechanism of the AAC remain unknown.

In this work we use coevolution analysis to obtain insights into the potential conformational dynamics of the MCF. By this approach, a multiple sequence alignment of homologous proteins was used to identify evolutionarily coupled residue pairs,

thereby indicating the pairwise spatial proximity of sites throughout the transporter. Paired residue interactions determined by coevolution analysis were largely consistent with both the solved AAC c-state and m-state structures. Comparison of the distances of the predicted coevolving pairs when mapped to the AAC structures, allowed us to identify important evolutionary constraints of both conformational states. *Ab initio* structural modeling based on our coevolution analyses provided insights into a model of AAC where both the matrix and cytosolic gates are closed. Our analysis also identifies important intra- and inter- module interactions of the transmembrane domain of the protein important for formation of the m-state. Our results therefore provide novel insights into important structural interactions applicable to all members of the MCF. Additionally, this analysis provides a unique insight into the strength of coevolution analysis to not only predict the structure of a protein but to inform conformational dynamics and interactions as well.

2.2 Introduction

2.2.1 Coevolution Analysis

The premise of coevolution analysis is that pairwise residue contacts are predicted based on strong evolutionary changes between physically interacting amino acid side chains (183-188). This is accomplished by generating a multiple sequence alignment (MSA) of the protein of interest for global statistical analysis to determine residue pairs that experience correlated changes over evolutionary time. In recent years, there have been vast improvements in the accuracy with which these contacts are predicted by using a global statistics model that is able to separate direct from indirect correlations (189-191). Multiple approaches to determine these correlations

have been proposed, including Generative REgularized ModeLs of proteINs (GRMELIN) (192), Direct Coupling Analysis (193) and Protein Sparse InverseCOVariance (194). GREMLIN, when optimized for contact predictions, has been shown to be most accurate (195,196). Another method, RaptorX Contact uses all three methods as input features to train a coevolutional neural network that is able to return a more accurate result for *ab initio* protein folding (197,198). Both the GREMLIN server and the RaptorX Contact server have been shown to accurately predict coevolving residues and structures of membrane proteins (197,199).

2.2.2 Investigating Structural Dynamics with Coevolution Analysis

Not only can coevolution analysis be used to predict protein structures but there is also precedent for the use of coevolution analysis to gain insights into the structural dynamics of proteins with multiple conformers. In 2012, Hopf et al. showed that coevolution analysis could be used to determine different conformational states of the transport protein, glycerol-3-phosphate transporter (GlpT). GlpT belongs to the facilitator superfamily that moves substrate across the cell membrane by alternating between two conformations, one open to the cytosol and the other open to the extracellular space. The structure of GlpT has been solved in the cytosol-open conformation and comparison of the structure to the coevolving pairs revealed a number of long distance residue pairs that informed contacts of the alternative conformation. Additionally, when just the coevolution analysis was used to predict the structure of GlpT, an intermediate conformation was observed, consistent with the idea that both conformations are evolutionarily constrained. These results suggest that coevolution analysis prediction

can illuminate precise conformational changes upon substrate binding and transport (200).

2.2.3 Application of Coevolution Analysis for Understanding the Mitochondrial Carrier Family

The MCF provides an excellent case study for understanding how coevolution can provide information about structural states, transport mechanisms, and interactions that drive structural change. Coevolution analysis predicts residues pairs that are proximal in the protein structure and it is generally thought that residue pairs predicted to be distal in a structure indicate false positive hits. However, when we consider a protein that is structurally dynamic, residue pairs that are distal in one structure could inform residue pairs proximal in an alternate conformational state (Fig 2.1). The known structural information of the AAC provides a unique opportunity to demonstrate not only the strength of coevolution analysis to predict protein structure but also to inform possible intermediate models and provide more detail about the transport mechanism and necessary interactions of the conformational states of the AAC. Due to the conserved structural elements of the MCF and the evolutionary nature of coevolution analysis these results, while mapped to the AAC, could also be used to infer information about other members of the MCF and therefore are widely applicable.

2.3 Materials and Methods

2.3.1 GREMLIN

GREMLIN (192,196) captures the conservation and coevolution of a MSA. This method optimizes the learning procedure by using a pseudo-likelihood approach, which

has been shown to be more accurate than previous methods that utilize DCA (195). GREMLIN uses a global statistical model that assigns a probability to each amino acid. The probability takes into account the position specific amino acid propensities and amino acid coupling between positions as determined by the MSA and maximizes the regularize pseudo-likelihood learning procedure. The strength of the measured coevolution within a protein family is indicative of physically interaction residue side chains.

2.3.2 Alignment Generation

The alignment generation was performed through GREMLIN. GREMLIN takes an input sequence and uses HMM-HMM-based lightning-fast iterative sequence search (HHblits: 2016_02) to find homologous sequences (201,202) and create a MSA. Only hits with an E-value of $1E^{-60}$ or above were considered. The alignment was processed so that all sequences had at least 90% coverage (no more than 10% gaps) and that all columns within the MSA that had more than 25% gaps were removed. Finally, the MSA was filtered so that no two sequences had greater than 90% sequence identity. GREMLIN was run using AAC input sequence from *S. cerevisiae* (318 residues).

2.3.3 GREMLIN BETA

GREMLIN BETA (199) evaluates the significance of a match between predicted contacts and a protein structure. The total GREMLIN score for a structure is determined for residues with a sequence separation of six or greater using P(contact). This expected GREMLIN contact score is compared to the observed score from the coevolution analysis. The ratio of observed contact score to expected contact score is

referred to as “Rc”. Rc values range from 0.7 to 1.2 for coevolution results compared to native protein structure contacts and 0 to 0.3 for contact maps paired with random structures. GREMLIN BETA also provides distances between coevolving pairs when mapped to a protein structure where the distance is defined as the minimal distance between any two atoms of the residue pairs, excluding hydrogen residues. Throughout this paper, we refer to residues in close proximity as residues that are within a distance of 5Å of one another.

2.3.4 Protein Coordinate Structures

The Protein Data Bank (203) was used to obtain the structures of AAC used for analysis by GREMLIN BETA. PDB ID: 4C9G (*S. cerevisiae*) was used to assess the fit of GREMLIN coevolution results to the c-state structure and PDB: 6GCI (*T. thermophilia)* was used to assess the fit of the results to the m-state structure.

2.3.5 RaptorX Contact

RaptorX Contact (197,198) predicts residue contacts using a deep learning method and integrates a variety of information including sequence conservation, residue coevolution, and contact occurrence patterns. An ultra-deep neural network is formed through the use of two deep residual networks. The first, takes into account the 1-dimensional sequential features. The second takes into account pairwise information and potential. RaptorX Contact takes these predicted contacts and converts them to distance restraints in order to perform *ab initio* protein folding without any templates or force fields. This approach allows protein structure prediction based solely on coevolution.

2.4 Results and Discussion

We performed coevolution analysis for the MCF and mapped residues onto the AAC as a strategy for elucidating the potential structural dynamics of this protein family. For this evaluation we compared our results to the known c-state structure of AAC (PDB: 4C9G) and the m-state structure of AAC (PDB: 6GCI). The key premise of this study is that residue pairs that show strong evolutionary correlation, but are not proximal in both conformational structures, are the evolutionary drivers of the transport mechanism of the AAC.

2.4.1 Comparison of AAC Coevolution Analysis with Known Structures

A MSA was generated with the sequence of *S. cerevisiae* AAC as the query sequence, using a stringency (E) value of 10^{-60}. The alignment generated by GREMLIN resulted in 7,326 unique sequences. A critical metric for the quality of a MSA for coevolution analysis is the ratio of the number of independent sequences to the total length of the protein (200). A sequence to length ratio of ~3 is generally accepted as a minimal value for accurate coevolution analysis (200,204). Our MSA had a sequence to length ratio of 25, underscoring the robustness of the information used for this analysis. With this dataset, we used the GREMLIN server to evaluate pairwise evolutionary residue correlations for the MCF. The analyses predicted a total of 433 coevolving pairs.

The results of coevolution analysis can be depicted as contact maps, which are two-dimensional representations of protein coordinates of identified coevolving residue pairs. The contact map generated from our analysis is mapped onto the sequence of *S. cerevisiae* AAC (Fig 2.2). Within a contact map, points along the diagonal denote

residue pairs that are proximal in the primary sequence. Such pairs typically exist within the same secondary structural element (e.g, separated by one turn along an a-helical axis). Unless indicated otherwise, we will define such primary contacts as sites that are separated by no more than three residues. Pairings between residues separated by longer distances on the primary sequence can be used to evaluate tertiary contacts within a protein. For a-helical bundle proteins, stretches of coevolving pairs perpendicular to the diagonal indicate antiparallel helical packing, whereas those stretches of coevolving pairs that are parallel to the diagonal indicate parallel helical packing. This contact map shows multiple clusters of residue pairs that are perpendicular to the diagonal, which results from the extensive known antiparallel interactions within and between the three modules of the proteins of the MCF.

We first evaluated coevolving residues with respect to known AAC structures. 398 residue pairs could be mapped to the c-state structure and 429 residues pairs could be mapped to the m-state structure. Table A.1. indicates the number of coevolving pairs sorted by structural region as identified by the 4C9G structure. Using the GREMLIN BETA server, all residue couplings were compared to both the c-state (PDB 4C9G) and the m-state (PDB 6GCI) structures (Fig 2.3). The contact maps compare residue pairs predicted to be coevolving (blue dots) with the residues known to be proximal within a given structure (grey dots) (Fig 2.3A). The overlap between the two coordinate sets for both structures show reasonably good agreement. Among all identified coevolving pairs within the c-state and m-state structures of the AAC, 70% and 68% were proximal (within 5Å) in the respective structures. Considering only coevolving pairs separated by more than three residues in the primary sequence, 65% and 90% were within 5Å and

10Å, respectively, when mapped to the c-state structure, and 62% and 90% were within 5Å and 10Å, respectively, when mapped to the m-state structure (Fig 2.3B).

Another way to assess the fit of our predicted contacts to known structures, is the Rc value, which is the ratio of the observed total GREMLIN server scores to the total scores expected from the protein structures (199). By this metric, our coevolution analyses compared well with structural information for the c-state (Rc = 0.771) and m-state (Rc = 0.796). Taken together, these results show good agreement between our coevolution analysis and the available crystal structures.

2.4.2 Comparison of the Predicted Coevolving Pair Distances when Mapped to the AAC

For both the c-state and m-state structures, several residue pairs were identified as distal (Fig 2.3B, red connections). This subpopulation of long-distance residue pairs, which constitutes approximately 10% of the coevolving residue pairs for each structure, are potential conformer-specific evolutionary pairings. We believe these pairs are potentially the interactions that drive the transition of the AAC between the two conformational states. To determine if these long-distance residue pairs are the consistent with an alternating access transition mechanism, we mapped the intermediate and long distance pairs (Fig 2.3B) and the residues involved in those pairs (Fig 2.4) on both the c-state and m-state structures. In agreement with the accepted alternating access mechanism, these residue pairs are located exclusively inside the cavity of both structures with long-distance pairs oriented toward the IMS- and matrix-facing sides of the transporter in the c- and m-state conformers, respectively.

Comparing the predicted coevolving pairs with the known interactions of both the c-state and m-state allowed us to identify the residue pairs proximal in both structures or in just one conformational state. When the predicted pairs were mapped, 395 coevolving pairs could be compared between the c-state and m-state structures. When considering the distance of the pairs, pairs were sorted into one of three ranges. The pairs were considered in contact if they were less then 5Å apart, pairs were of intermediate distance if they were separated between 5Å and 10Å, and long distance if they were greater than 10Å apart.

Of these 395 predicted pairs, 78% did not have a significant change in distance (i.e. if the coevolving pair was less than 5Å apart when mapped to the c-state structure it was less than 5Å apart when mapped to the m-state structure). This high percentage indicates that coevolution analysis predicts the overall conserved core structure of the AAC between the c-state and m-state that has been discussed previously (159).

Of the 22% of coevolving pairs that did have a significant change in distance 54% of coevolving pairs are proximal in the c-state but distal in the m-state and 46% are proximal in the m-state but distal in the c-state. Comparing the coevolving pairs with the known contacts of the c- and m-state we can identify the residue pairs that are only in contact in one state (Fig 2.5, green ovals). Comparison of the predicted pairs to the two different sets of known contacts of the c- and m-state structures (Fig 2.5, top half of the diagonal) shows that in each structure there are three regions of proximal coevolving pairs that become distal in the alternate conformation. We believe the residues proximal in only one conformational state of the AAC indicate the evolutionary constraints that stabilize either the c- or m-state. The identified residue interactions are listed in Figure

2.5. Analysis of these residues shows the regions proximal in only the c-state are residues of the odd-numbered helices, whereas the m-state only proximal regions are residues on the even-numbered helices. This is consistent with a transition mechanism in which the odd-numbered helices come together and form the matrix gate in the c-state and the even-numbered helices come together and form the cytosolic gate in the m-state. Lending even more credibility to this analysis is the fact that we also have identified residue interactions involved in the salt bridge motifs responsible for the formation of the c- and m-state gates (Fig 2.5, bolded residue interactions). Together with the structures of the AAC we were able to verify the coevolution analysis by identifying coevolving pairs proximal in only one conformational state that have been shown to be necessary for formation of one state or the other. We extend these finds and indicate a number of other evolutionary important interactions for the c- or m-state conformation.

2.4.3 AAC Model Generation using Coevolution Analysis

With the verification of the GREMLIN server coevolution analysis we used RaptorX Contact to predict the structure of the AAC based solely on the predicted coevolving pairs (Fig 2.6). This model represents a structure of AAC that attempts to satisfy all predicted coevolving pairs; therefore, it likely captures critical interactions of both the c- and m-states. When compared to the GREMLIN server analysis, the Rc value of the predicted structure was 0.732 indicating good agreement between the model and predicted coevolving pairs. This is evident when the predicted coevolving pairs were mapped to the model, with 62% of the coevolving pairs within a distance of 5Å and 95% of all predicted pairs being within a distance of 10Å (Fig 2.7). The 5% of

coevolving pairs that exist at a distance greater than 10Å appear to shift away from either end of the cavity towards the center of the protein and towards the area of substrate binding (Fig 2.7B, 2.8). Analysis of the contact map compared to the known hits of this model shows that the interactions important for AAC transition identified in Figure 2.5 are all within 5Å of each other in this predicted model (Fig 2.7A, green ovals). Together this suggests that this model attempts to satisfy the greatest number of predicted coevolving pairs and does so by minimizing the number of long distance contacts. By minimizing the number of long distance coevolving pairs, closure of both gates of the cavity occurs which is evident upon comparison of the model to the c-state (Fig 2.9A) and m-state (Fig 2.9B) structures.

2.4.3.1 Comparison of the Cavity Area

To compare the closure of the cavity at either end of the protein we determined the area of the cavity in the c-state, m-state, and RaptorX Contact model (Fig 2.10). Not surprisingly, at the cytosolic side of the protein the c-state structure had the largest cavity area and at the matrix side of the protein, the m-state structure had the largest cavity area. In theory, if the RaptorX Contact model represented an intermediate conformation of the AAC the area of the cavity would be somewhere in between the c-state and m-state structures along the entirety of the cavity. What we see is that at both sides of the model the cavity area is smaller than either the matrix side of the c-state structure or cytosolic side of the m-state structure. This overall decrease in the area of the cavity suggests the predicted model may instead represent an occluded state of the AAC that satisfies all the most important evolutionary constrained residues resulting in an even tighter closure of both cavity gates.

2.4.3.2 Comparison of the Matrix and Cytosolic Gate Residue Distances

To compare the closure of both the matrix and cytosolic gates, we mapped the coevolving pairs that involved residues of the signature motifs, Px[D/E]xx[K/R] and [Y/F][D/E]xx[K/R], and the glutamine and tyrosine brace residues.

Six coevolving pairs are found to include residues between the matrix gate motif residues or the glutamine brace residues (Fig 2.11). In agreement with the alternating access mechanism, in the c-state structure most of the coevolving pairs are closer than 5Å apart whereas in the model and m-state structures these residues get further apart with the majority being separated by either 5-10Å or greater than 10Å apart (Table A.2A).

Four coevolving pairs were identified between the cytosolic gate residues and the tyrosine braces (Fig 2.12). Again in agreement with the alternating access mechanism, in the m-state and model structures, where the cytosolic gate is closed, the residues of these pairs are closer than 5Å apart. In the c-state structure, where this gate is open, the residue pairs are all greater than 10Å apart (Table A.2B).

Together this coevolution analysis accurately predicted some of the most important interactions for the closure of the matrix and cytosolic gates. By mapping these pairs onto the AAC structures we can confirm how the distance of these pairs changes depending on weather the gates are opened or closed. Additionally, when comparing the distances of these pairs on the model to the solved structures we confirm the closed nature of both gates and that the necessary interactions were maintained and predicted by the RaptorX Contact model.

In evaluating this RaptorX Contact model in comparison to the known c- and m-state structures we have identified several takeaways. First, structural prediction by coevolution analysis should be interpreted with caution because by trying to satisfy all the coevolving pairs, this model does not provide an accurate representation of either conformational state. Instead, by maximizing the number of proximal contacts and minimizing distal contacts we observe a kind of “hybrid”, which in the case of AAC results in the closure of both cavity gates. While we cannot say this model represents an actually intermediate state, presumably the conformational landscape between the c- and m-states may be occupied with many conformers that are intermediate in nature and possibly resemble this model. It is also possible that due to the higher number of short and intermediate distance contacts that this model represents a much lower energy conformer than the c- and m-state structures which have been locked in a conformational state through the use of an inhibitor.

2.4.4 Understanding Intramodular Interactions through Coevolution Analysis

2.4.4.1 Agreement of Intramodular Interactions with Predicted GxxxG and πxxxπ Motifs

Previous research has shown that GxxxG motifs on the odd-numbered helices and πxxxπ (small – xxx- small) motifs on the even-numbered helices are important for the tight packing of the transmembrane helices on the cytosolic side of the protein (159). In this analysis we see that the residues of these motifs coevolve with one another both within a motif (i.e., separated by three residues within the same transmembrane segment), and between two intramodule motifs (Table 2.1). Along with these coevolving pairs the small residues of these motifs are involved in a number of

other coevolving pairs. These pairs highlight interactions between the small residues of the πxxxπ motifs and larger bulky residues on the odd-numbered helices. The coevolution analysis suggests that these interactions are important for stabilization of the m-state and closure of the cytosolic side of the protein. With the inward movement of the even-numbered helices the bulky residues on the odd-numbered helices are able to slide out towards the lipid bilayer and interact with the πxxxπ residues in a knobs into holes like packing that results in multiple intramodule interactions and stabilization of the m-state intramodule helices (Fig 2.13).

Module one shows additional intramodule interactions involving the GxxxG motif residues that suggests bulky residues on the adjacent transmembrane helix are important for tight intramodule packing in the m-state. The GxxxG motif is located on helix one and coevolves with bulky residues of helix two that, in the m-state, shift toward the center of the cavity. This allows the glycine residues of H1 to shift to the outside of the protein while still allowing tight intramodule interaction and m-state stabilization (Fig 2.13).

Together this analysis suggests that the importance of the small residues of the GxxxG and πxxxπ motifs extend past just simply allowing tight packing of the intramodule helices. These residues are also important for stabilizing the closure of the cavity by interacting with bulky intramodule residues of adjacent helices that provide multiple points of intramodule interaction.

2.4.4.2 Analogous Intramodule Coevolving Pairs

Since the AAC is made up of three tandem repeats (modules one, two, and three) we hypothesize that the most important coevolving interacting pairs for the transition mechanism would be predicted in all three of the modules. Analysis of the coevolving pairs occurring within each module revealed about 100 pairs. Of these 100 pairs, 50 indicated tertiary contacts. To understand how these identified coevolving residues compared within each module, the three modules were aligned and the pairs were compared. We believe this analysis highlights the residues most important for stabilizing the intramodule interactions of the AAC. These residues are highlighted on the c-state (Fig 2.14A), RaptorX Contact model (Fig 2.14B), and m-state (Fig 2.14C) structures and identify specific faces of both the transmembrane helices and matrix helices.

We identified eight analogous coevolving pairs occurring in the transmembrane region of each module. The residues appearing on the even-numbered helices include residues of the πxxxπ motif along with other residues that fall along the same helical face. Mapping these connections on the c-state structure (module one shown for clarity) suggests the dynamic nature of this intramodular interaction (Fig 2.15, left). Based on the predicted coevolving pairs alone, it suggests the N-terminal end of H1 is proximal to the C-terminal end of H2, however in the c-state structure this is not the case (Fig 2.15, left). Looking at the distances between these pairs when mapped to the model (Fig 2.15, middle) and the m-state structure (2.15, right) there is a decrease in the distance of the pairs that were distal in the c-state resulting in all identified pairs being less than 5Å apart in the m-state structure. Module two and three follow a similar pattern suggesting a symmetrical movement of the cytosolic ends of the intramodule

transmembrane helices towards each other in the m-state across all modules (Fig A.1, A.2, Table A.3). The significance of these predicted connections is two-fold. First, we have identified the intramodule helical faces are important for stabilization of the transmembrane region of the protein and how they interact depends on the conformational state of the AAC, which was not apparent from conservation analysis alone. Second, due to the available structures of the AAC we can validate that these helical faces are important for the intramodule interactions of the transmembrane domain.

2.4.5 Evaluation of Coevolution Analysis

2.4.5.1 Analysis of Non-Coevolving Residues

While a high percentage of AAC residues were identified in this coevolution analysis we found 22 residues that do not coevolve with any other residues. Presumably these residues were not identified because they are highly conserved and therefore not changing over time, or the residues are not making any structurally important tertiary contacts and therefore are not subject to evolutionary constraint imposed by other residues. Analysis of the identity of these residues shows approximately half are located in predicted loops, connecting structural elements, and therefore they are not subject to coevolution-detected constraints. Conservation analysis identified the other half as highly conserved. Here, we identified the proline residues of the Px[D/E]xx[K/R] motif of the odd-numbered transmembrane helices and six glycine residues. Due to the role of the proline residue in matrix gate formation and providing the necessary kink of the odd-numbered helices it is of no surprise these residues do not appear in the coevolution analysis. When analyzing the six glycine

residues, we identify a new motif potentially important for CL binding. First, three of the six identified glycine residues (G89, G192, G287) (Fig 2.16, blue residues), located at the amino-terminal end of the even-numbered helices, are part of one of the CL binding motifs, [YWF][KR]G (178,205). CL has been shown to promote AAC stability (206) and as such these glycine residues and the glycine residues of the second CL binding motif, [Y/F]xG (Fig 2.16, red residues), located at the amino-terminal end of the matrix helices, are believed to provide structural flexibility to the region that results in the exposure of positively charged amide groups (157) with which, the negatively charged CL can interact. These glycine residues (G89, G192, G287) are highly conserved due to their necessary role in CL binding making it not surprising they do not appear in the coevolution analysis. The other three glycine residues (G82, G185, G280) (Fig 2.16, green residues) are also highly conserved, but have not been identified as part of any previously identified motif. Together these six glycine residues occur in a GxxxxxxG motif present in each module. GxxxxxxG motifs are commonly found in transport proteins (207,208), and are thought to enhance helical packing. However, this does not seem to be the case considering the location of this motif. Therefore, we hypothesize that the GxxxxxxG motif together with the additional CL binding motif ([Y/F]xG) form a glycine triangle that is important for CL binding. Previous MD simulations lend credence to this hypothesis by identifying both glycine residues of the GxxxxxxG motif of module three as CL interactive residues (181). Perhaps these newly identified glycine residues aid in the stabilization of CL binding during the transition between conformational states and help contribute the necessary flexibility of this region to allow continued CL binding and optimal AAC function.

2.4.5.2 Analysis of Highly Coevolving Residues

As previously suggested there could be a link between the strength of the evolutionary constraint on a residue and the functional importance of that residue (i.e. the residue is involved in interfacial regions, binding sites, or are potentially involved in conformational changes (200)). To see if this might be the case for our analysis of the MCF we assessed the strength of evolutionary constraint on a residue by summing the GREMLIN BETA server score values over all coevolving pairs involving that residue (Fig 2.17A). The residues that scored within the top 10% are found in two general areas of the AAC, the intermodule helical faces of the even-numbered transmembrane domains and the matrix helices.

Five residues out of the top 10% of highly coevolving residues occur in the interfacial region between helices of adjacent modules (H2:H3, H4:H5, H6:H1) (Fig 2.17B, Table A.4). By identifying the pairs these highly coevolving residues are involved in we determined that these residues and the residues they coevolve with highlight the helical face important for intermodule interactions upon closure of the m-state. Previous research has suggested that small residues of the odd-numbered helices are important for intermodule interaction however, we identify here that these highly coevolving residues (I111, A136, L140, L214, L308) are important for the stabilization of the intermodule interactions necessary for m-state formation.

The majority of the highly coevolving residues are found in the matrix helices, (Fig 2.17C, Table A.5-A.7) there could be multiple explanations for this. First, these matrix helices could be imperative to the structure of AAC due to the high number of i, i+3 and i, i+4 connections made by the highly coevolving residue pairs. Second,

residues within these matrix helices may be more highly subject to tertiary packing constraints owing to their solvent accessibility. Transmembrane helices are constrained to assume an a-helical structure in order to satisfy backbone hydrogen bonding in the low dielectric membrane environment. By contrast, the matrix helices of AAC are exposed to water and therefore do not have the same secondary structural constraints imposed by the local dielectric environment. The high scoring nature of the evolutionary couplings in this region could reflect the importance of maintaining the α-helical shape of the matrix helices, which are also stabilized by tertiary contacts with the rest of the protein. It also could be that the high coevolving nature of this region is a result of the necessity of CL binding in this vicinity. The identified GxxxxxxG motif occurs in the same region as these highly coevolving residues. It could be that the conserved nature of these glycine residues has put greater evolutionary pressure on the nearby matrix helices resulting in exceedingly high coevolution as a mechanism to adapt more readily to any evolutionary changes necessary to maintain CL binding.

2.5 Conclusions

The AAC is not only an essential mitochondrial IM transporter but it is also the most abundant and widely studied protein of the MCF informing most of the current knowledge of the structures of this protein family. The MCF provides a unique opportunity to assess the strength of coevolution analysis in understanding the conformational dynamics of proteins with multiple states and to determine protein structure. Although there is structural knowledge of both the c-state and m-state of the AAC, they are locked in these states by chemical inhibitors, because of this; the dynamics of the transition of the AAC remains an open question. To address this

question, we take a new approach to analyze the MCF using coevolution and mapping our results onto the AAC.

In this study, we predict global structural changes that occur during the transition of the AAC from the c-state to the m-state and investigate the residue-residue communication across the whole protein and within specific modules of the AAC. We identify a number of novel intra- and inter- module interactions of the transmembrane domain that stabilize the formation of the m-state. The intramodule interactions involved the previously determined small residues of the transmembrane helices and have been thought to allow sliding of the transmembrane helices across one another and allow for tight helical packing. In this study, we have identified these small residues also interact with bulky residues on the adjacent transmembrane helices within a module that mimic knobs into holes helical packing and aid in the stabilization of the m-state. Additionally, we determined novel intermodule interactions of the transmembrane domain in the m-state that are not evident from the solved structures.

Due to the available structures of the AAC we can identify that 80% of coevolving pairs have no significant change in distance between the two conformational states confirming previous results that the core structure of the AAC remains unchanged. The remaining 20% of coevolving pairs, those that either got significantly closer together or further apart, were used to identify interactions most important for the transition of the AAC. In this subpopulation of coevolving pairs, we identify not only experimentally determined interactions of the residues of the matrix and cytosolic gates but identify new interactions that, in the future, should be experimentally tested for functional significance. This analysis also substantiates that the c- and m-state structures are not

just stabilized by the inhibitors found in the available crystal structures, but that there is a clear pattern of coevolution that suggests inter- and intra- module interactions that promote tertiary packing that are inherent in the protein sequence.

Our analysis is also a great case study for understanding the strength of coevolution analysis. It is important to remember the only input for this analysis was a MSA and that these results predict a number of already experimentally proven interactions and the core structure of the AAC. The determined RaptorX Contact model correctly predicts this core structure and satisfies the greatest number of long distance coevolving pairs seen in either the c-state or m-state structures. Coevolution analysis predicts that the residues of the cytosolic and matrix gates are in proximity to one another, validating the alternating access mechanism and demonstrating the strength of coevolution to predict proximal residue pairs of multiple conformational states. Finally, due to the sequence and structural similarities of the MCF proteins the insights determined here are widely applicable and could be mapped onto any protein of the MCF.

2.6 Figures

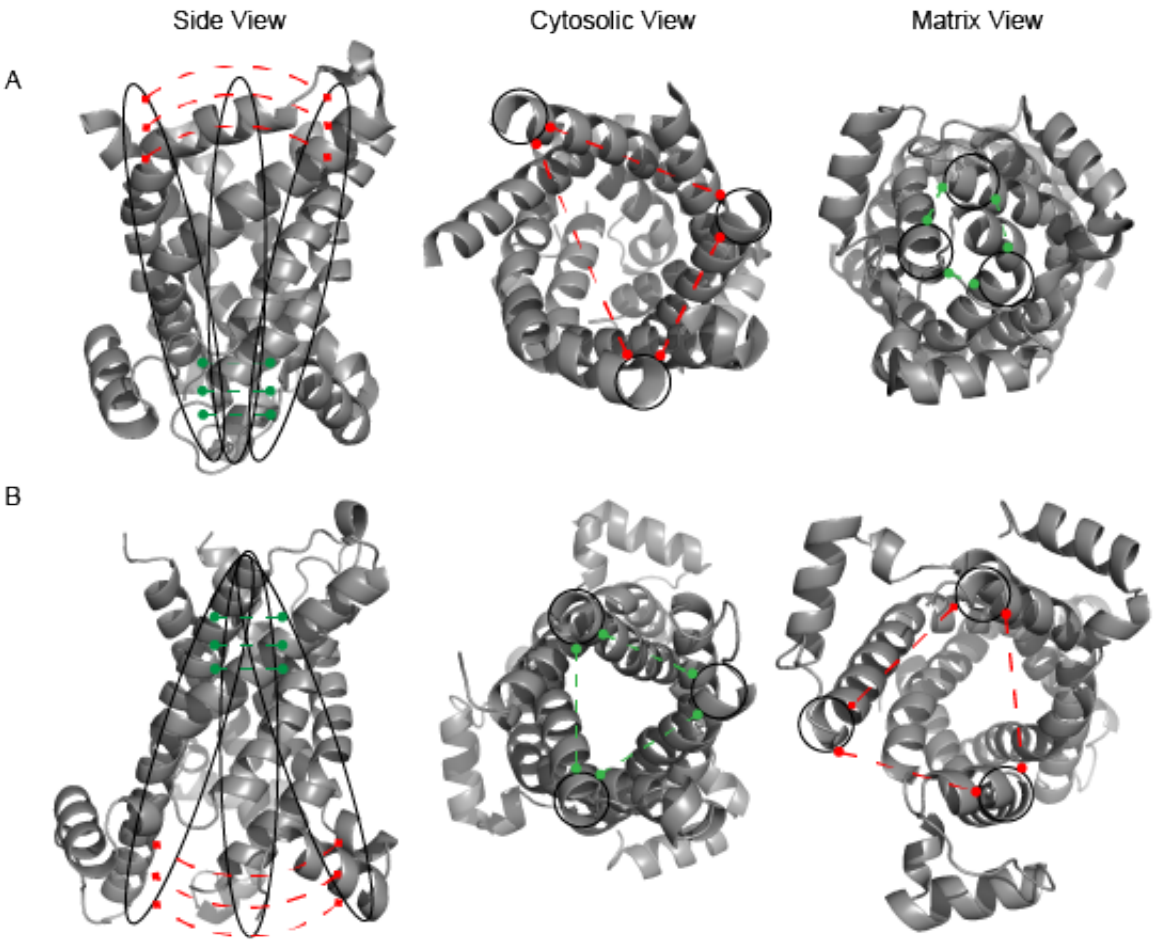


Figure 2.1. Predicted short and long distance coevolution patterns. Principle behind using long distance coevolving sites to inform alternate conformations of a protein. (A) Identified coevolving pairs on the c-state AAC structure may be proximal (here, toward the matrix side shown as green connections) or distal (here, toward the cytosolic side shown as red connections). (B) The accepted alternating access mechanism results in distal sites moving closer together and proximal sites moving further apart in the m-state AAC structure.

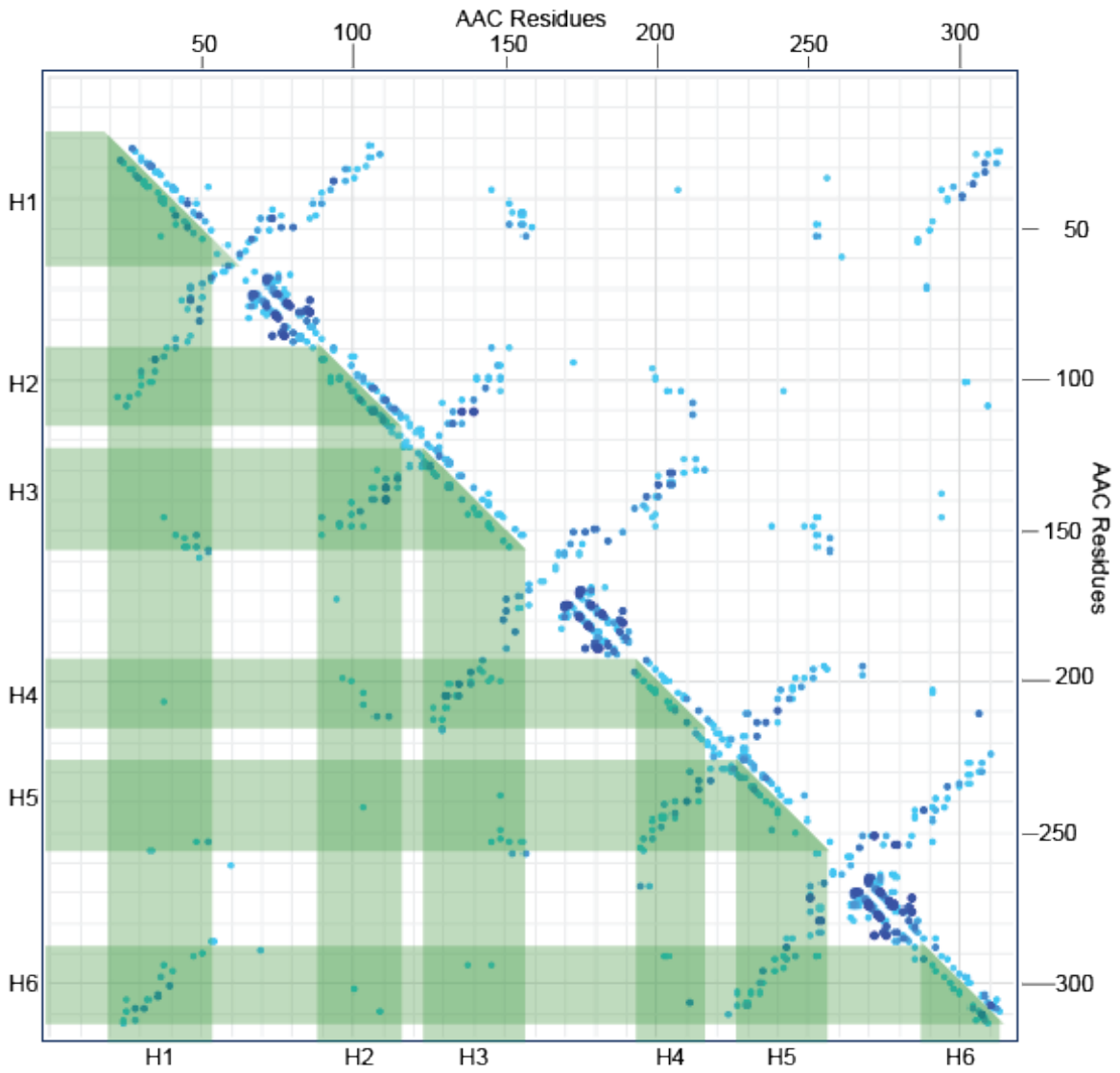


Figure 2.2. Contact map of coevolution results. Axes represent indices of the yeast AAC primary sequence and green boxes indicate boundaries of transmembrane helices H1-H6. Blue points represent a pair of identified coevolving residues, the darker the blue the stronger the coevolution score.

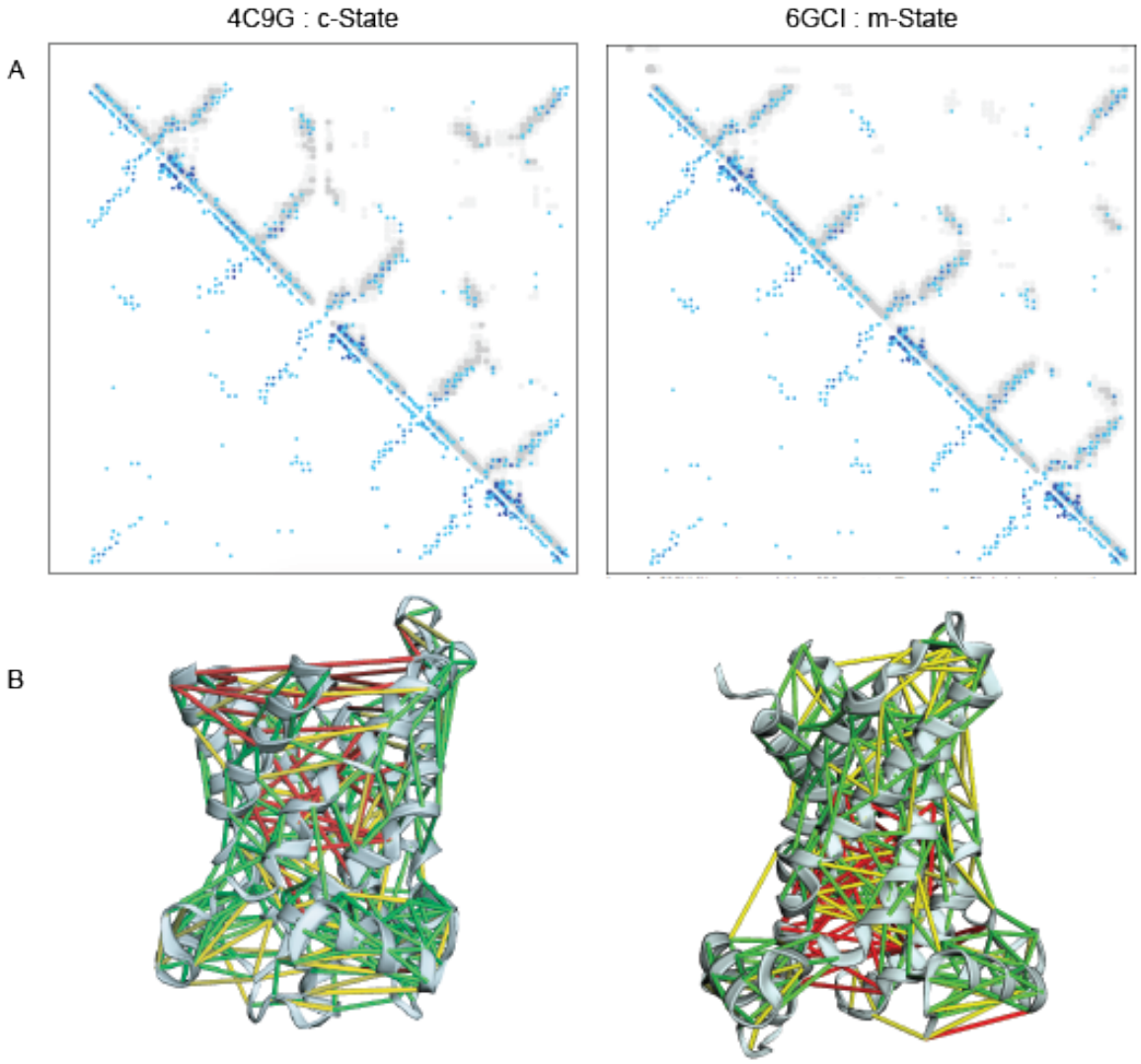


Figure 2.3. Contact map comparison to known AAC structures. (A) Comparison of coevolving pairs and pairwise proximity in the c-state structure, *left*, and the m-state structure, *right*. Blue points represent sites of identified coevolving residues, the darker the blue the stronger the coevolution score. Grey points represent proximal residues in the c-state AAC structure (PBD 4C9G) and the m-state (PDB 6GCI) with contact distances indicated as dark grey (nearest heavy atoms 5Å apart) and light grey (nearest heavy atoms separated by 5-10Å).

(B) Residue pairs are mapped on to the c-state, *left*, or m-state, *right*, structures.

(green, <5Å apart; yellow, 5-10Å apart; red, >10Å apart)

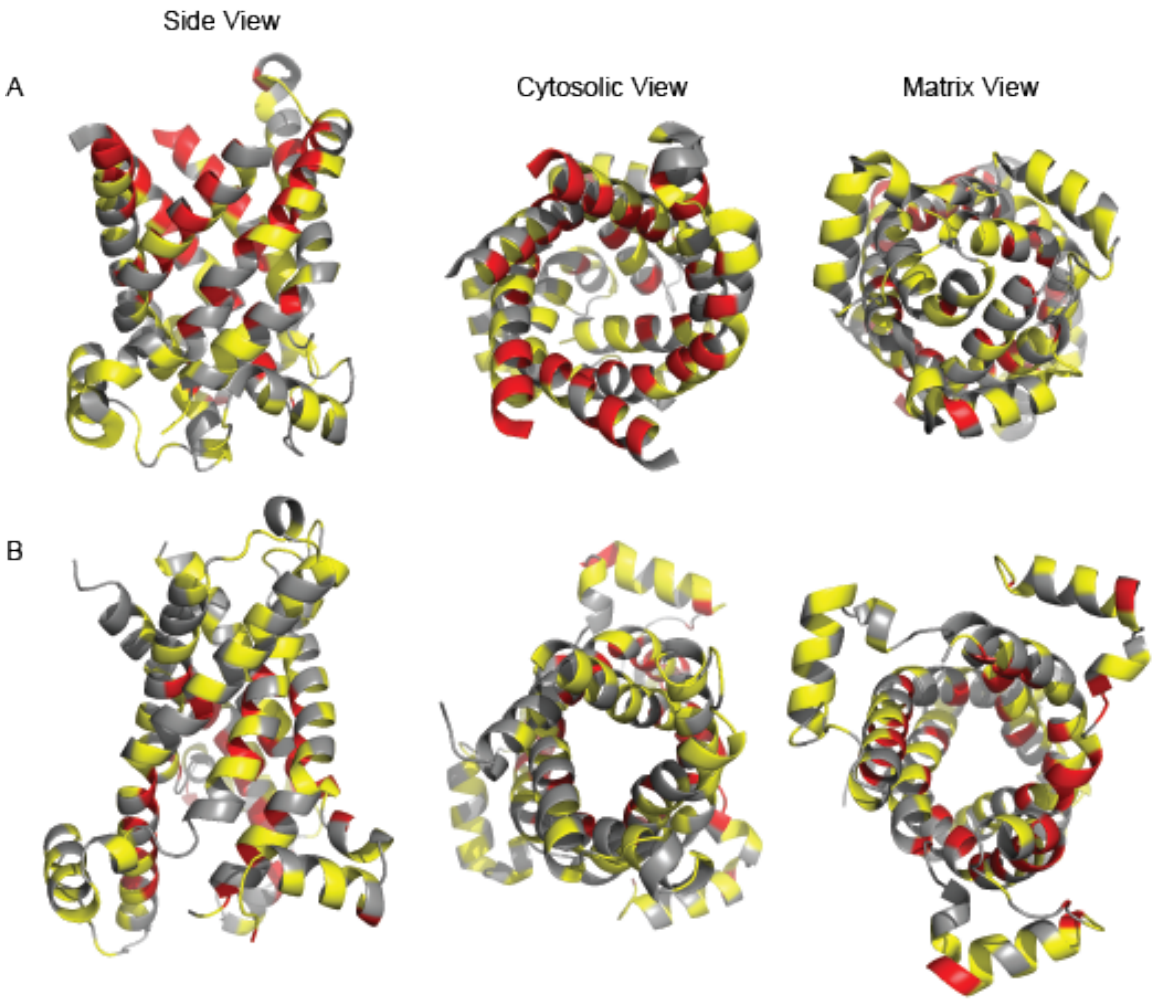


Figure 2.4. Residues involved in intermediate and long distance coevolving pairs. Ribbon structures shown with residues involved in coevolving pairs of an intermediate distance (5-10Å) colored yellow and long distance (>10Å) colored in red. *Left*, side view; *middle*, cytosolic view; *right*, matrix view. (A) c-state structure (PDB 4C9G). (B) m-state structure (PBD 6GCI).

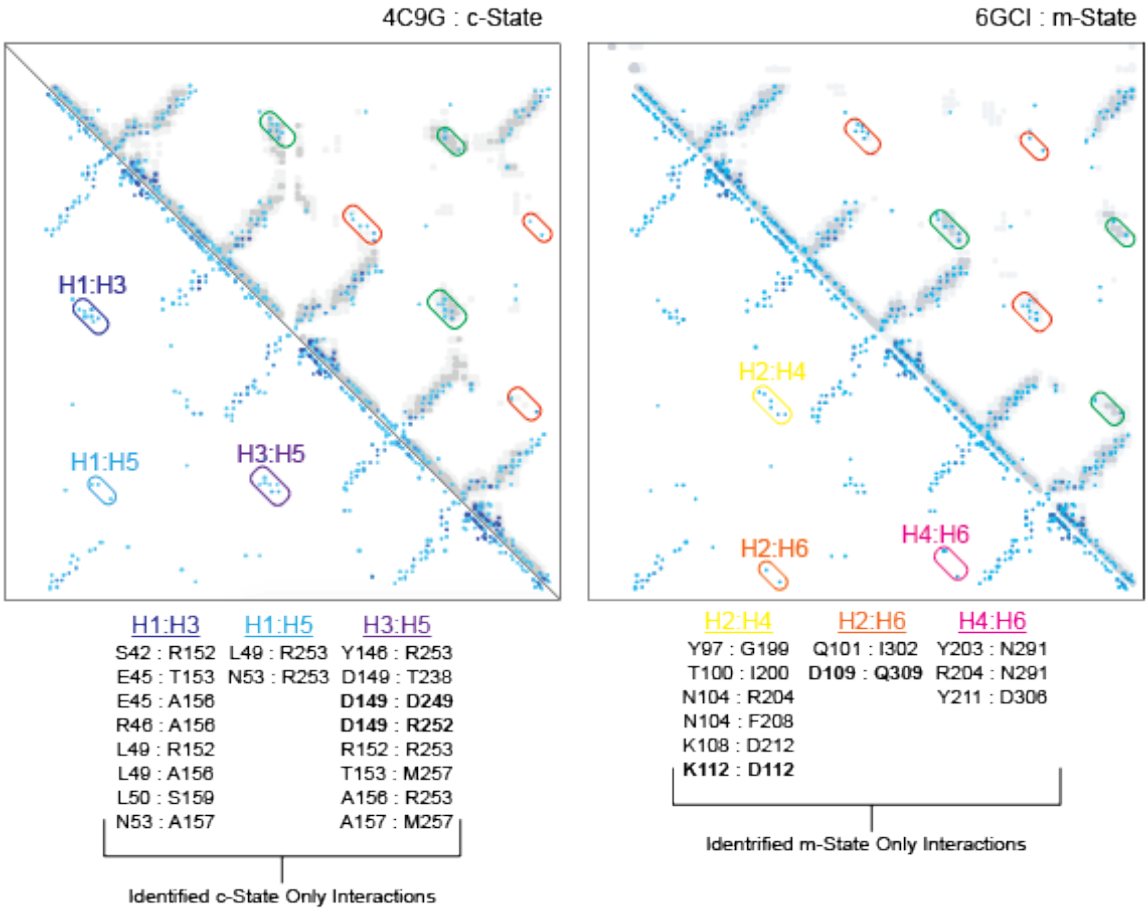


Figure 2.5. Coevolving pairs identified as proximal in only one conformational state. The top halves of the contact maps highlighted green ovals indicating proximal residue pairs and red ovals indicating distal residue pairs of *left*, c-state structure and *right*, m-state structure. The bottom half of the contact maps have color coded ovals that represent predicted residue pairs predicted to be important for the left, c-state formation and right, m-state formation. The residue pairs identified on the contact map are listed below, *S. cerevisiae* numbering is used.

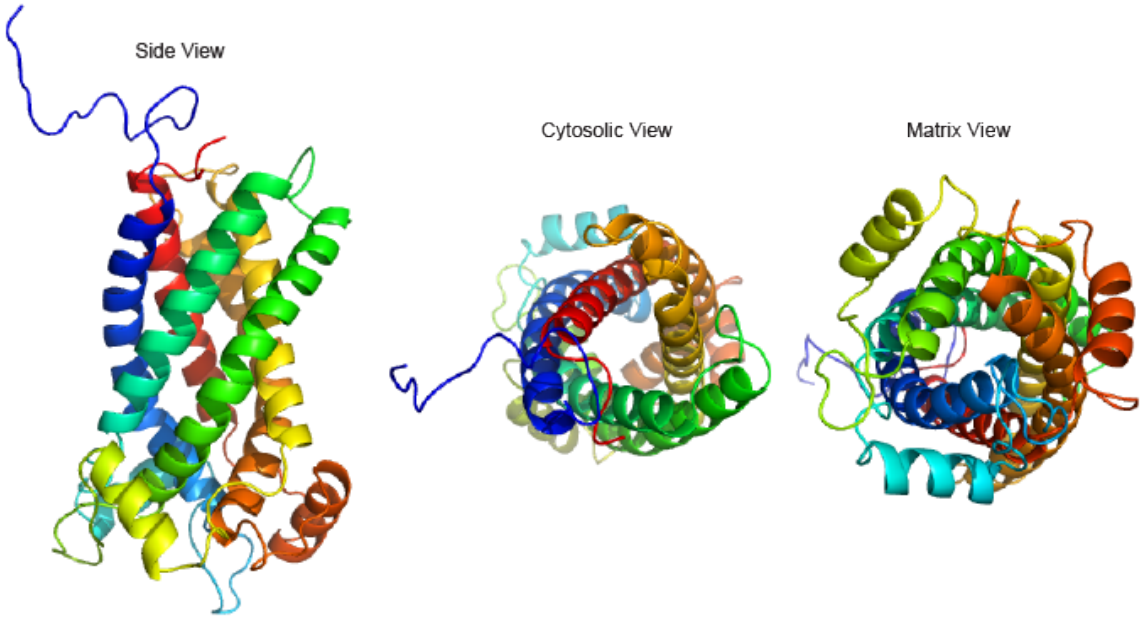


Figure 2.6. RaptorX Contact model. RaptorX Contact model of the AAC as predicted by RaptorX Contact. Backbone ribbon rainbow color-coded, N-terminus (blue) to C-terminus (red). *Left*, side view; *middle*, cytosolic view; *right*, matrix view.

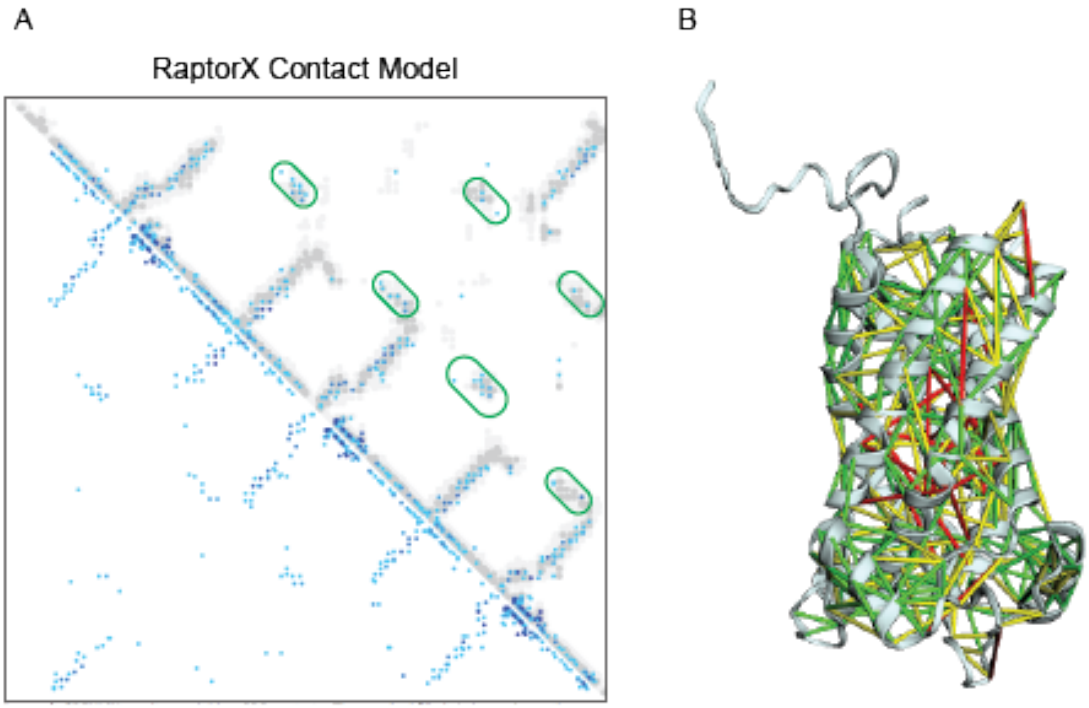


Figure 2.7. Comparison of RaptorX Contact model with predicted coevolving pairs. (A) Blue points represent sites of identified coevolving residues, the darker the blue the stronger the coevolution score. Grey points represent proximal residues in the RaptorX Contact model with contact distances indicated as dark grey (nearest heavy atoms 5Å apart) and light grey (nearest heavy atoms separated by 5-10Å). Ovals refer to the clusters of residue pairs identified in Figure 2.5. All residues pair clusters are now within a distance of 5Å. (B) Residue pairs are mapped on to the RaptorX Contact model (green, <5Å apart; yellow, 5-10Å apart; red, >10Å apart).

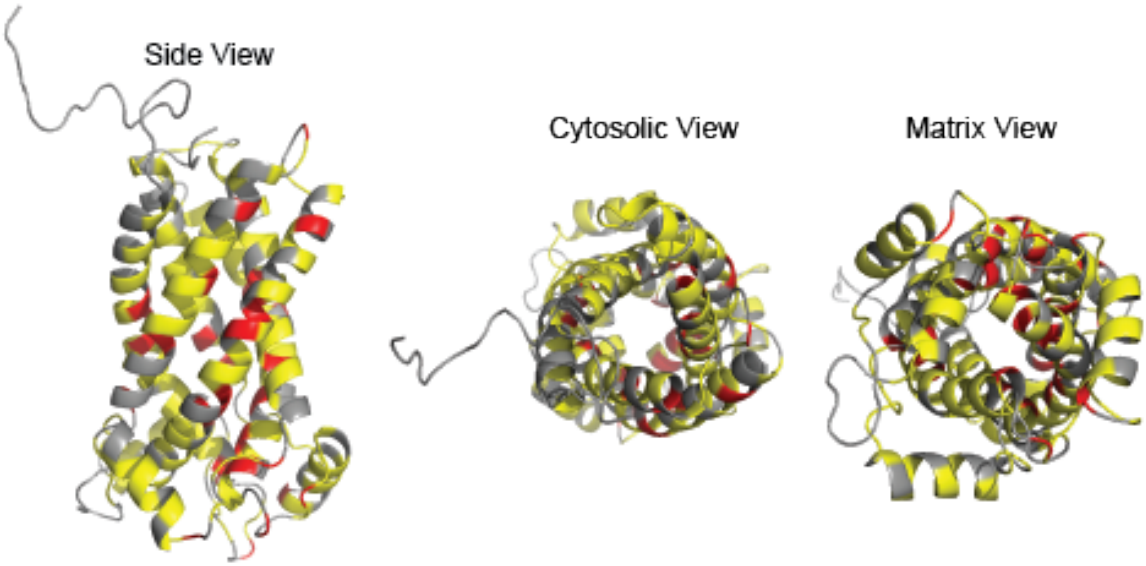


Figure 2.8. Residues involved in intermediate and long distance coevolving pairs in the RaptorX Contact model. Ribbon structure of the RaptorX Contact model colored to show residues involved in coevolving pairs of an intermediate distance (5-10Å) colored yellow and long distance (>10Å) colored in red. *Left*, side view; *middle*, cytosolic view; *right*, matrix view.

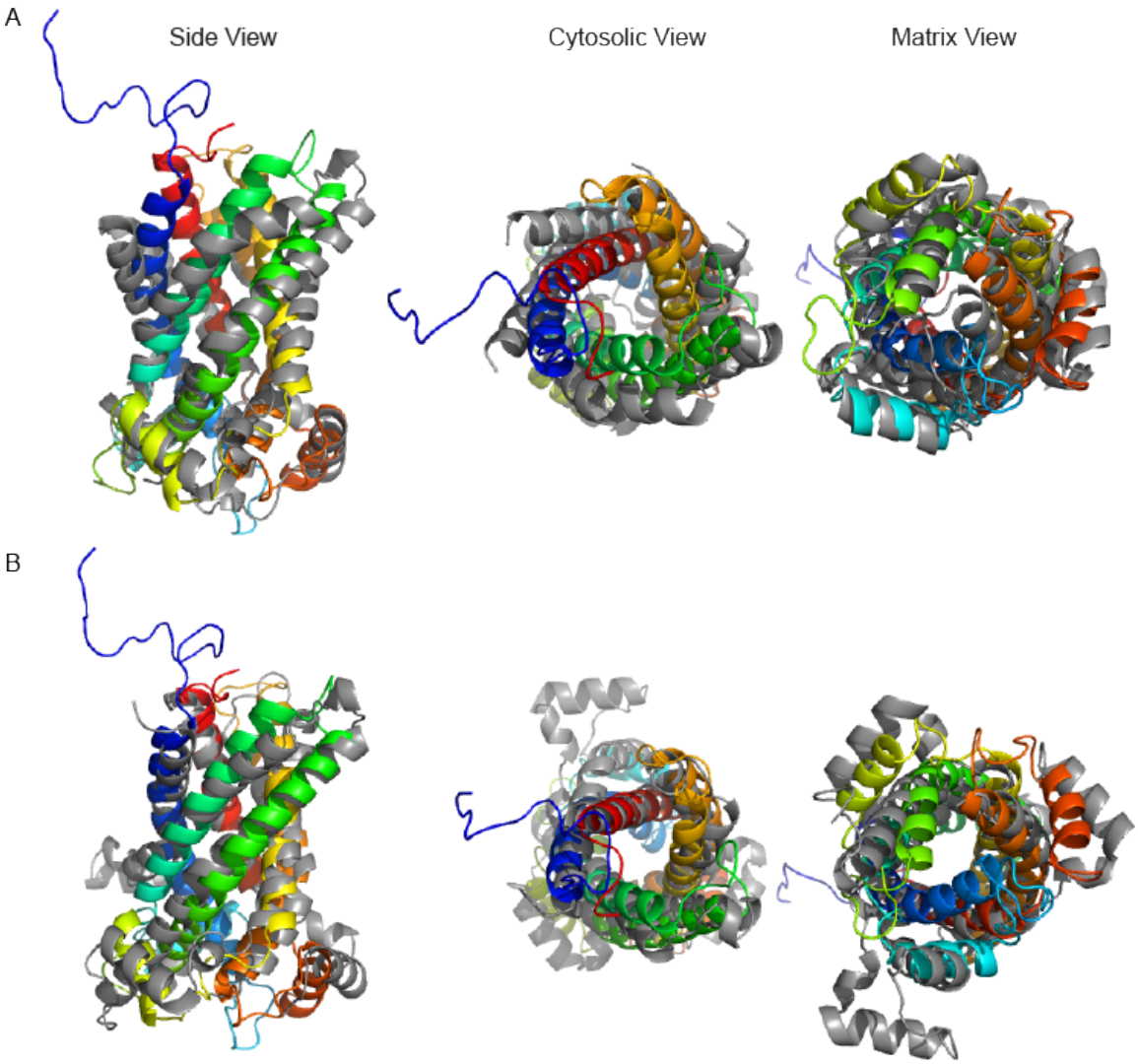


Figure 2.9. Comparison of RaptorX Contact model with known structures.

Model, colored, overlaid with known AAC structure for comparison. (A) 4C9G. (B) 6GCI. *Left*, side view. *Middle*, cytosolic view. *Left*, matrix view.

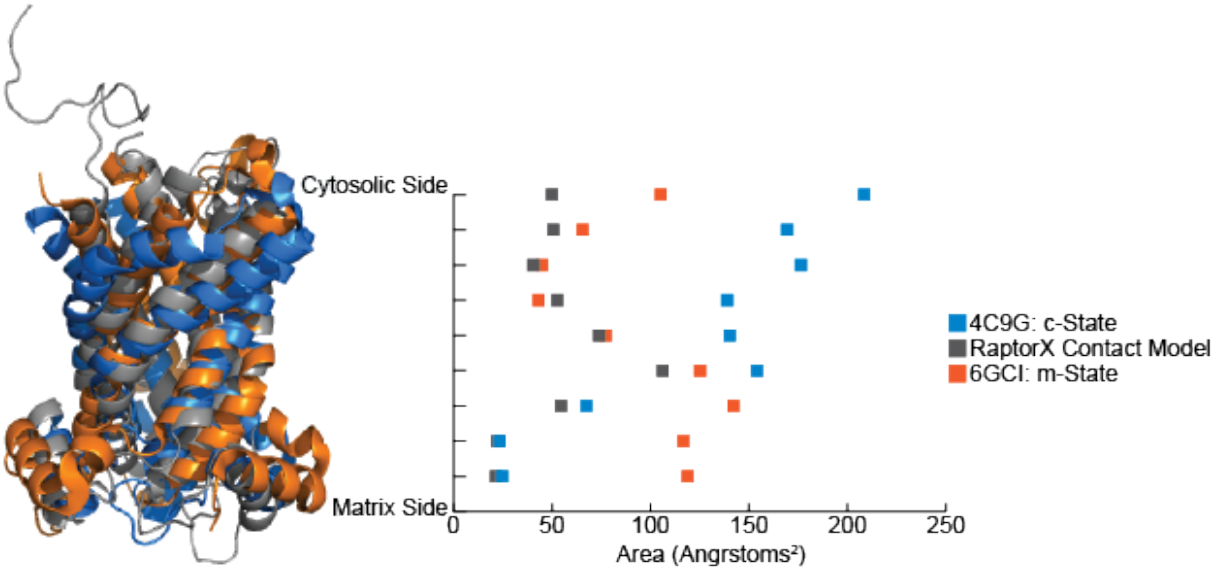


Figure 2.10. Cavity area analysis. The area of the cavity for the c-state, m-state, and model were analyzed. Channel-facing residues on the same transmembrane helical register were used to compute triangular areas defined by even-numbered helices on the cytosolic side of the protein and the odd-numbered helices on the matrix side of the protein for which the area ($Å^2$) was determined. *Left*, the three structures are aligned with the 4C9G structure in blue, the model in grey, and the 6GCI structure in orange. *Right*, the area analysis with the helical register in line with the structures to show the area of the cavity at different helical turns.

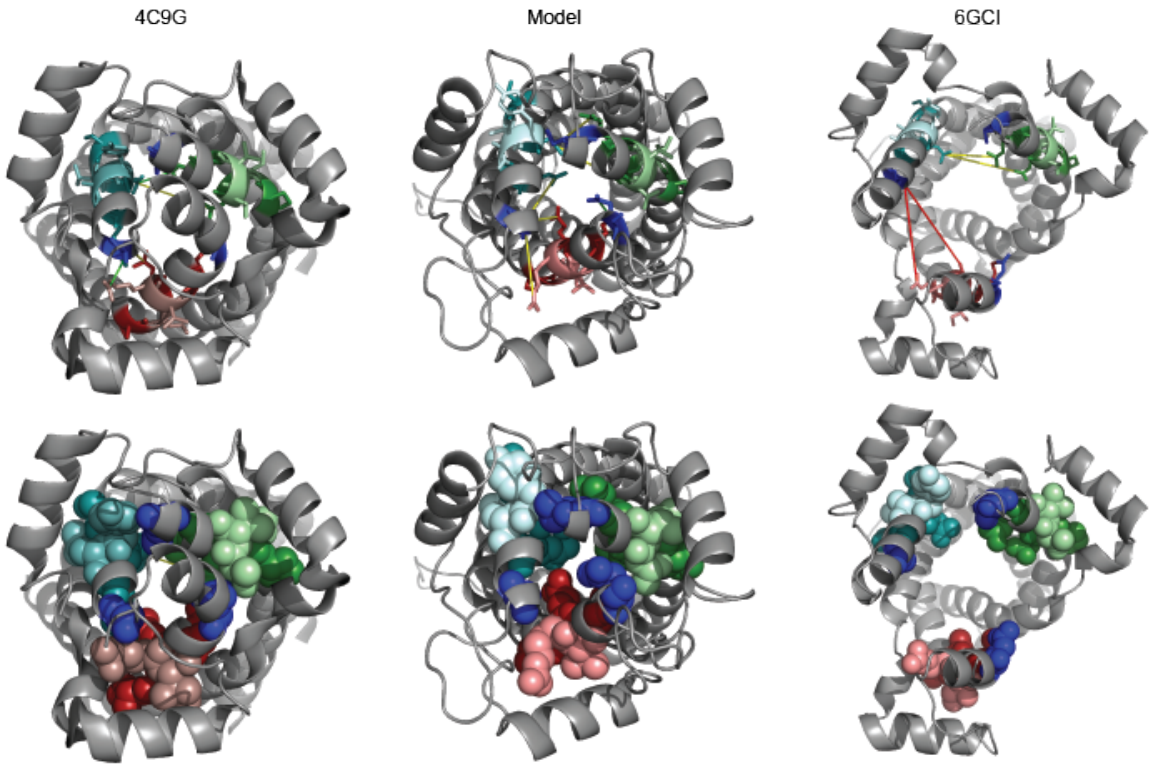


Figure 2.11. Comparison of the matrix gate. The c-state stabilizing motif Px[D/E]xx[K/R] highlighted in red (module 1), teal (module 2), and green (module 3). Glutamine brace residues shown in blue. *Left*, 4C9G; *Middle*, RaptorX Contact model; *Right*, 6GCI. Top panel, residues shown as sticks; bottom panel, residues shown as spheres.

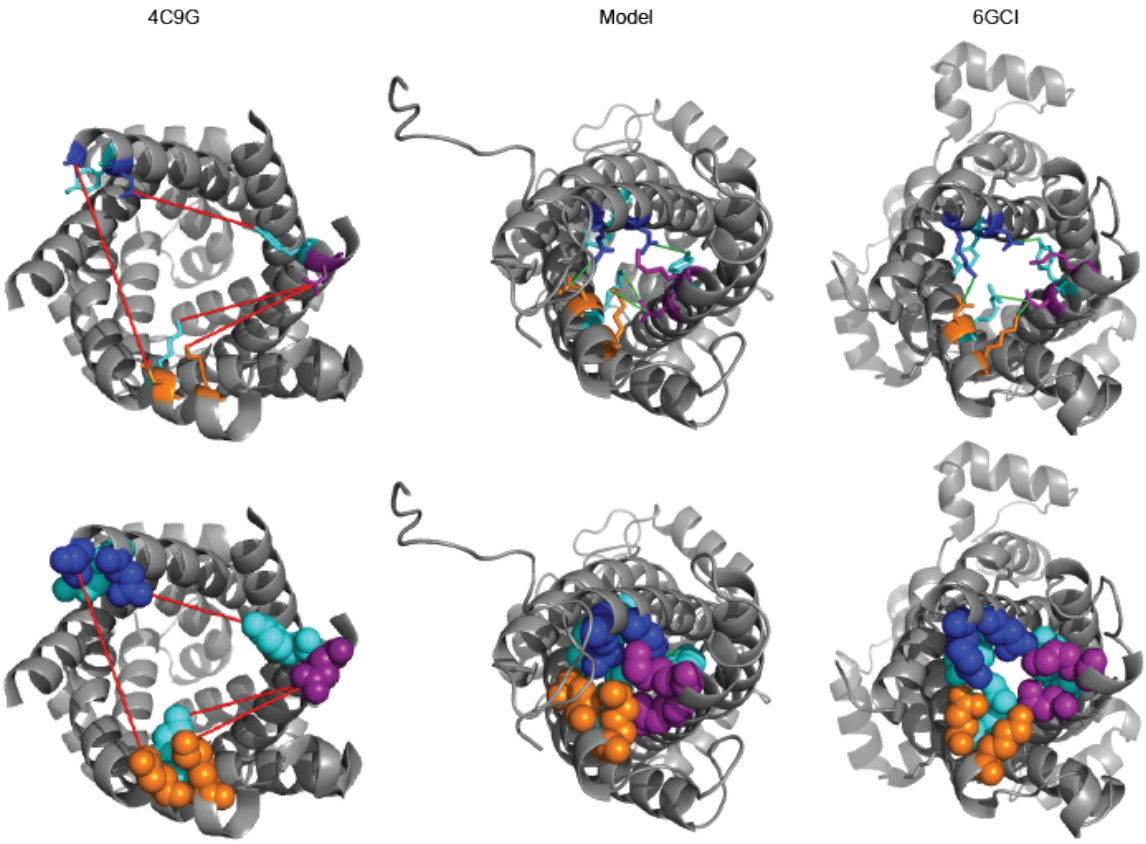


Figure 2.12. Comparison of the cytosolic gate. The m-state stabilizing motif [FY][DE]XX[RK] highlighted in orange (module 1), purple (module 2), and blue (module 3). Tyrosine brace residues shown in cyan. *Left*, 4C9G; *Middle*, RaptorX Contact model; *Right*, 6GCI. Top panel, residues shown as sticks; bottom panel, residues shown as spheres.

	1st Predicted Residue	2nd Predicted Residue	Distance Catagory 4C9G	Model	6GCI
Intra-Motif Interactions	Gly 30/22	Ala 34/26			
	Ala 102/94	Ala 106/98			
	Gly 134/127	Gly 205/198			
	Gly 205/198	Gly 209/202			
	Gly 234/227	Thr 238/231			
	Gly 234/227	Ser 303/296			
Intra-Module Interhelical Interactions	Gly 30/22	Phe 105/97			
	Ala 34/26	Gln 101/93			
	Gly 31/23	Ala 102/94			
	Phe 27/19	Ala 106/98			
	Leu 131/124	Gly 205/198			
	Gly 135/128	Gly 205/198			
	Phe 127/ Met 120	Gly 209/202			
	Leu 131/124	Gly 209/202			
	Phe 231/224	Ser 303/296			
	Gly 234/227	Gly 300/293			
	Gly 300/293	Ser 303/296			
	Pro 99/91	Ala 102/94			
	Ala 106/98	Lys 110/102			
	Leu 206/199	Gly 209/202			
	Thr 238/231	Ala 241/234			
	Gly 138/131	Val 145/138			
	Ala 34/26	Met 256/249			
	Gly 138/131	Arg 294/287			
	Asp 149/142	Thr 238/231			
	Thr 238/231	Asp 249/242			

Key

- GxxxG Motif
- πxxxπ Motif
- GxxxG Motif Interacting Residues
- πxxxπ Motif Interacting Residues
- <5Å Seperation
- 5-10Å Seperation
- >10Å Seperation

Table 2.1 Residue pairs of the GxxxG and πxxxπ motif interactions. Table listing all the coevolving pairs that involve the glycine or small residues of the GxxxG and πxxxπ motifs.

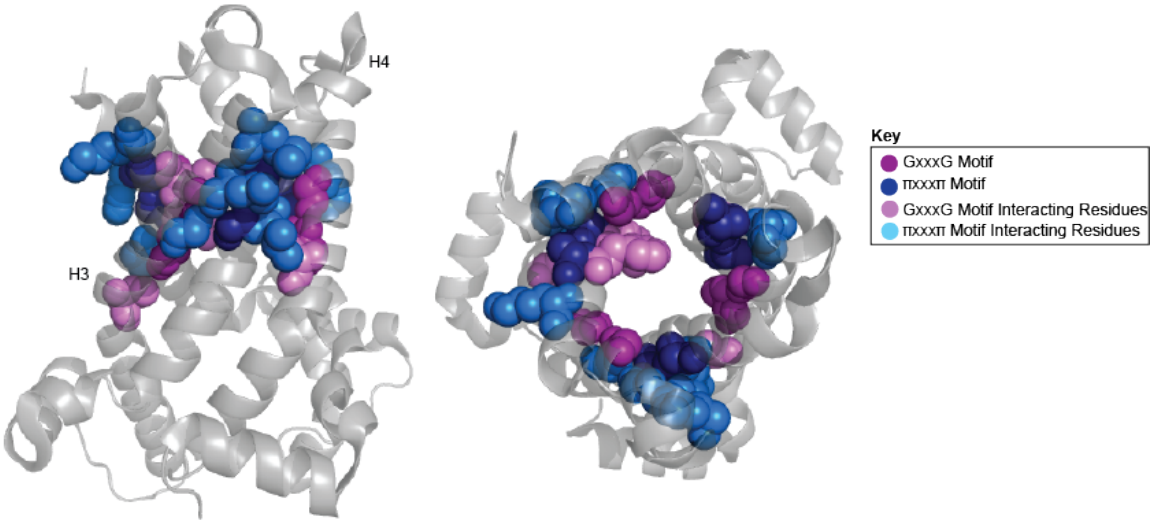


Figure 2.13. GxxxG and πxxxπ motif interactions. Spheres indicate residues on the 6GCI structure involved in these interactions and colored coded according to the key. *Left*, side view. *Right*, cytosolic view.

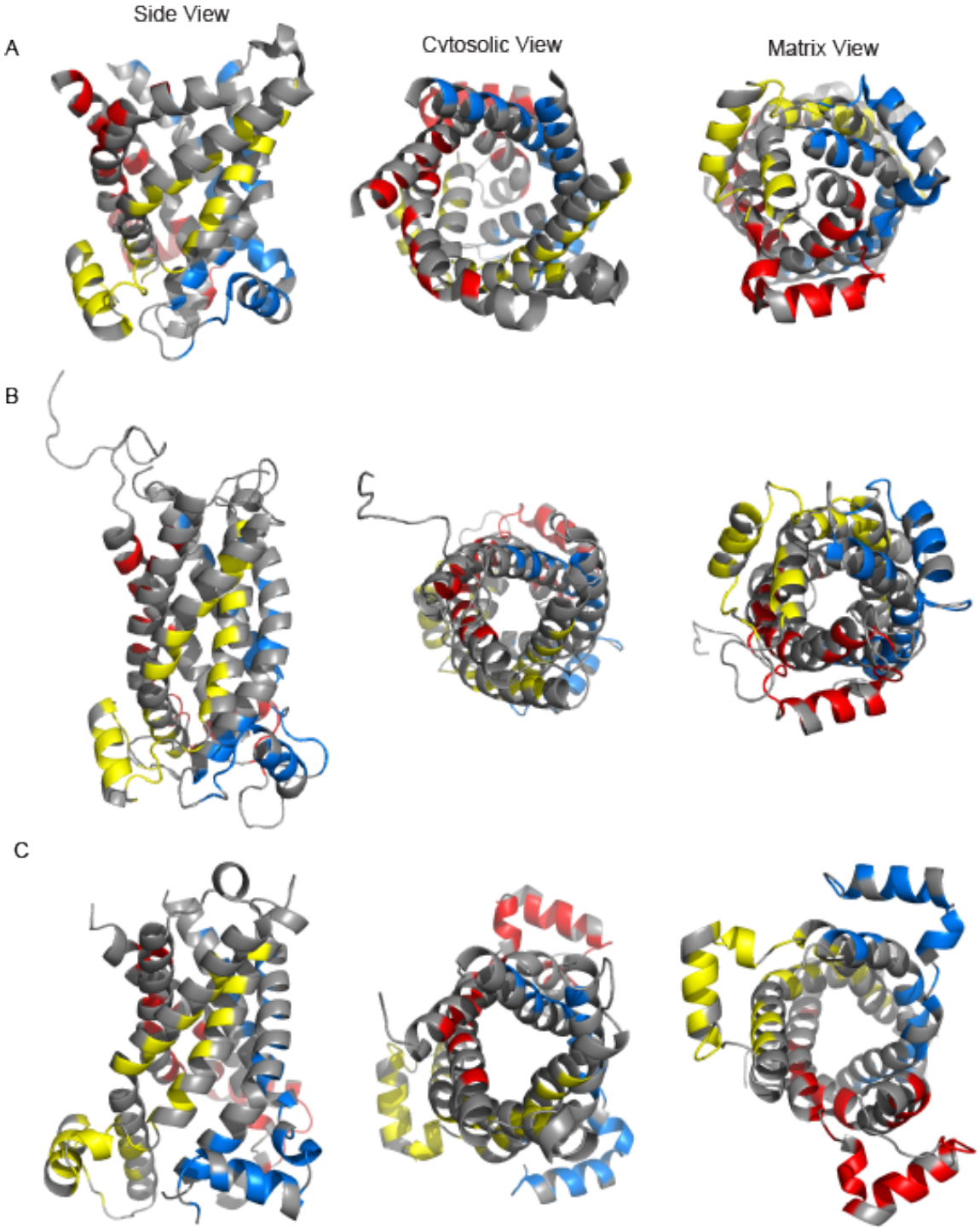


Figure 2.14. Intramodule analogous interhelical interactions. Intramodular coevolving pairs that occur in all three modules were identified. The residues

involved in these pairs are colored demonstrating the helical faces predicted to be evolutionarily important. Intramodule interactions represented in red (module 1), yellow (module 2), and blue (module 3) mapped onto (A) 4C9G. (B) RaptorX Contact Model. (C) 6GCI. *Left*, side view. *Middle*, cytosolic view. *Right*, matrix view.

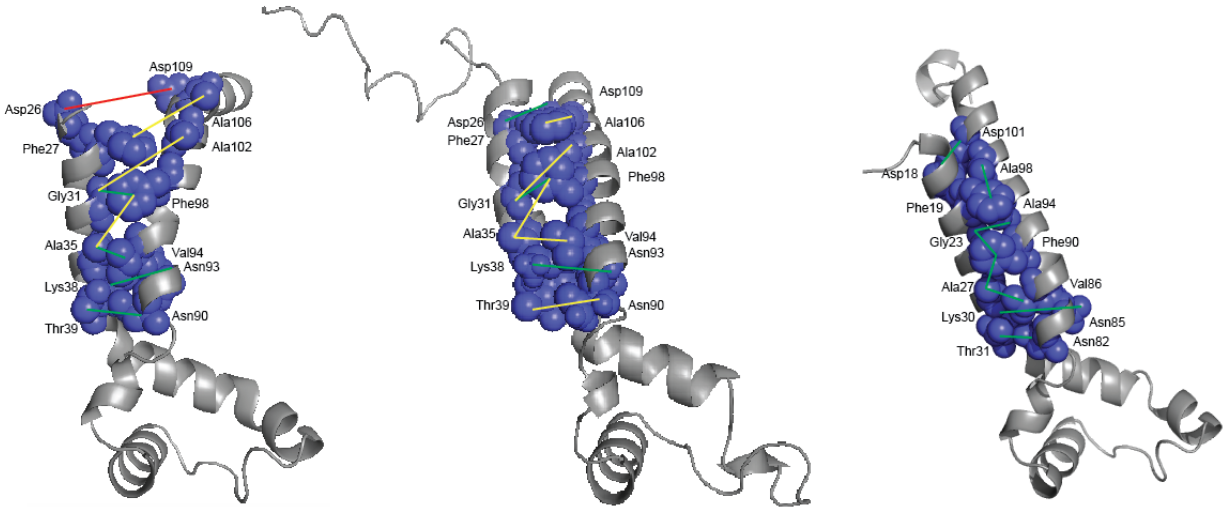


Figure 2.15. Intramodule transmembrane helical packing. The coevolution analysis identified residues in a classical i, i+4 pattern along one helical face of both H1 and H2 (blue spheres). These faces suggest the helical face important for intramodule interaction, module one shown. *Left*, 4C9G. *Middle*, RaptorX Contact Model. *Right*, 6GCI. (green, <5Å apart; yellow, 5-10Å apart; red, >10Å apart)

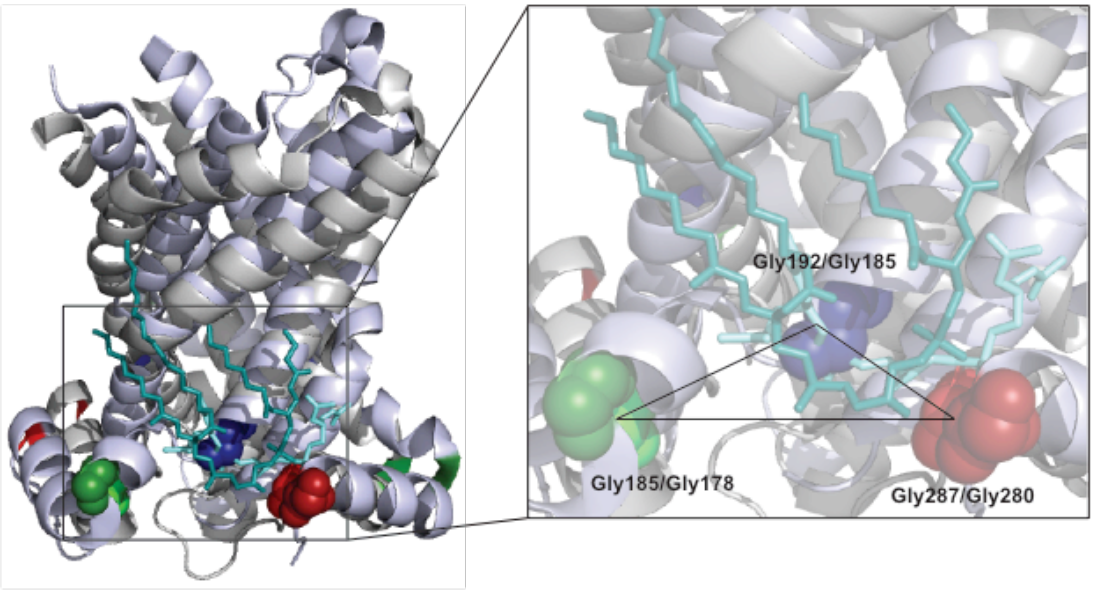


Figure 2.16. CL interactive glycine triangle. 4C9G (pale grey) and 6GCI (pale blue) were aligned. CL shown as sticks; pale teal indicates CL molecules resolved in the 4C9G structure; dark teal indicates CL molecules resolved in the 6GCI structure. Residues of the identified glycine triangle shown as spheres with blue and green residues highlighting the GxxxxxxG motif. Red spheres highlight the glycine residues of the additional CL binding motif. Insert, highlights residues of the CL binding site between modules 2 and 3 illustrating this structural feature.

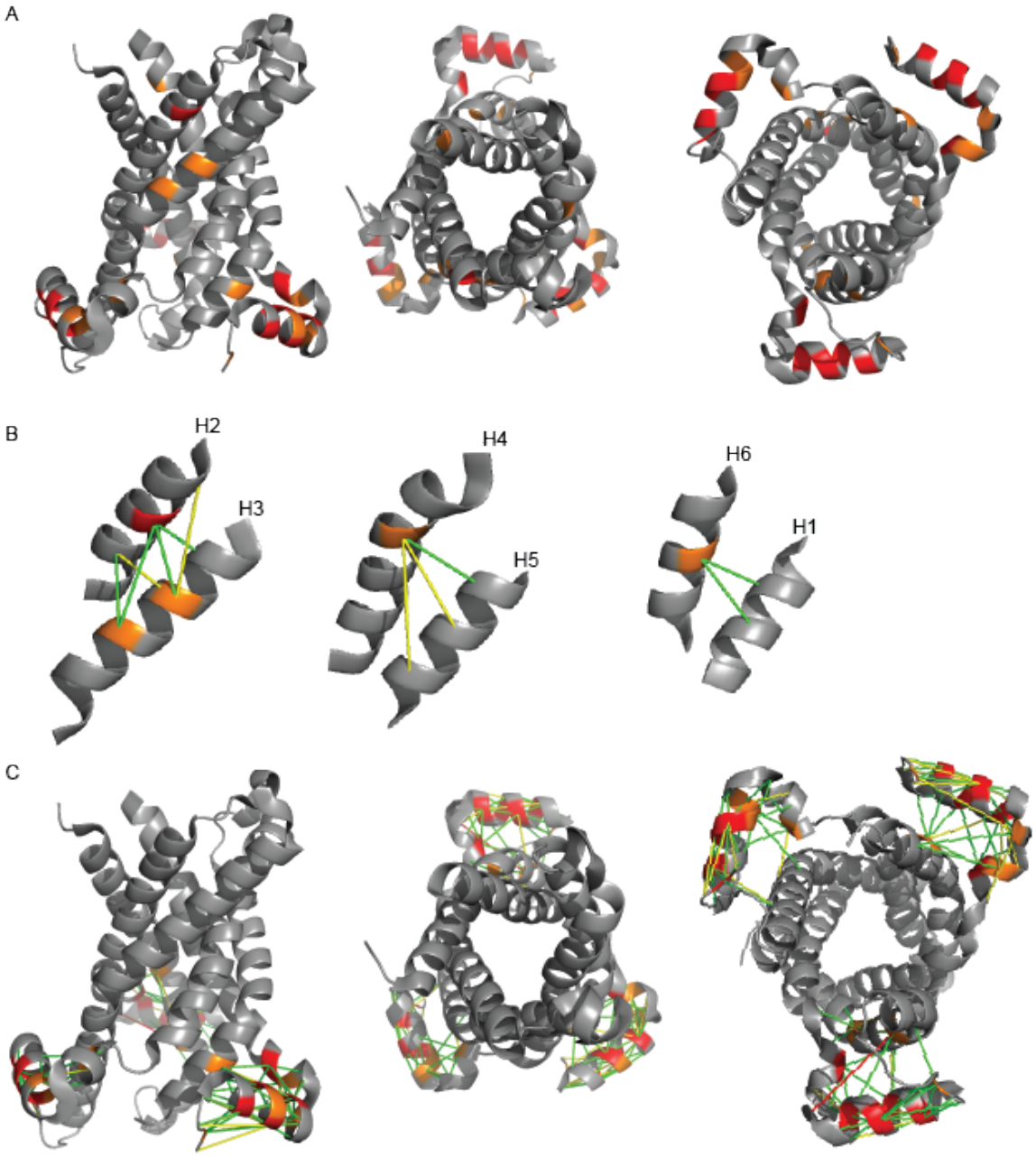


Figure 2.17. Highly coevolving residues. The GREMLIN Beta score for each residues pair was summed for each residue to determine possible functionally important residues. The top 5% of residues are shown in red and the top 6-10% of residues are shown in orange. (A) Residues colored and mapped on 6GCI.

Left, side view. *Middle*, cytosolic view. *Right*, matrix view. (B) Coevolving pairs shown to highlight the intermodule connections made by highly coevolving residues. (C) Coevolving pairs shown to highlight the dynamic structural nature of the matrix helices.

Chapter 3

Analysis of Tim23 Superfamily Proteins using Bioinformatics-Based Approaches

3.1 Abstract

The majority of mitochondrial proteins are synthesized in the cytosol and must be imported into the mitochondria. The TIM23 complex, located in the mitochondrial IM, is responsible for the import or integration of approximately 70% of all mitochondrial proteins. The channel domain of the TIM23 complex consists of membrane proteins Tim23 and Tim17. Previous studies have illuminated a number of key structural elements necessary for protein import and channel formation, however, there is currently no high-resolution structure of Tim23 or Tim17. Due to the challenges membrane proteins present when using traditional structure determination methods this study takes a new approach, coevolution analysis. Through the use of a MSA, coevolution analysis predicts residues proximal in the three-dimensional structure that can then inform distance restraints used for *ab initio* structure prediction. We identify coevolving residue pairs for the channel domain of Tim23 and using these results then perform *ab initio* modeling that informs the first structural model of the channel domain of Tim23. Analysis of the determined model and predicted structural contacts suggests the matrix helix of L1 is highly dynamic and potentially acts as a presequence-like structure that inserts into the transmembrane domain of Tim23 when the channel is not actively engaged by precursor protein. Additionally, we investigate channel formation by performing inter-protein coevolution analysis on Tim23 and Tim17 and identify residues of the matrix loops of the two proteins as coevolving. This study also investigates the N-terminal domain of Tim23 using conservation analysis to identify that the presence of the certain amino acids could be the important factor for OM interaction. This study lends novel insights into the structure of the channel domain of Tim23, potential

structural dynamics, and residues important for TIM23 subunit interaction and OM interaction that should be used to inform future experimental studies.

3.2 Introduction

3.2.1 Tim23 Structural Insights

Although there is no high-resolution structural information for Tim23 or Tim17, there are a number of key structural regions that have been identified as important for channel assembly and protein translocation (Fig 3.1). The N-terminal domain of Tim23 is unstructured and facilitates a number of interactions with the OM and other TIM23 complex subunits. A heptad leucine repeat motif mediates coiled-coil interactions that aid in the dimerization of Tim23 and interaction of Tim23 with Tim50, Tim21, Tom22, and presequence (102,209,210). Purified whole Tim23 and the channel domain alone have been shown to form 13-24Å wide channels (73). The channel domain of Tim23 is made up of four TMS with both the N- and C-termini located in the IMS (211). Tim23 also has an extended positively charged matrix loop (L1, predicted as 26 residues in *S. cerevisiae*) that links TMS1 and TMS2. GxxxG motifs mediate helical packing interactions between TMSs, thereby stabilizing both intra- and inter-molecule interactions among many membrane proteins (212). TMS1 and TMS2 contain conserved GxxxG motifs necessary for the structural integrity of the TIM23 complex (213). For example, mutations of G102, G149, G153 to leucine residues are lethal (Fig 3.1, Tim23, pink) and are thought to compromise the association of Tim23 with Tim17, Tim44, and Tim50 (213). Mutations of residues G112, G116, and G145 (Fig 3.1, Tim23, yellow) impair the interaction of Tim23 with Tim17, indicating that both TMS1 and TMS2 play a key role in the assembly of the core channel domain (213).

TMS2 contains an aqueous facing side that may define part of the translocation pore (214). Depolarization of the membrane causes a greater proportion of TMS2 to be exposed to the aqueous IMS region and interrupts the helical structure of TMS2 (215). This indicates that Tim23 is structurally dynamic and can assume different functional conformations, which could have distinct interactions with other TIM23 subunits. TMS2 has been shown to promote the cation selectivity of Tim23, and mutational analysis shows that N150A impairs yeast growth and decreases cation selectivity of the channel (216).

3.2.2 Tim17 Structural Insights

Tim17 has a similar structure to the Tim23 channel domain, including four predicted TMSs and the extended L1 located in the matrix (Fig 3.1, Tim17). Tim17 has four conserved cysteine residues found at positions 10, 77, 118, and 120 in *S. cerevisiae*; C10 and C77 form a disulfide bond that stabilizes the interaction between TMS1 and TMS2 (217) (Fig 3.1, Tim17, blue). Tim17, like Tim23, has a number of GxxxG motifs, and mutations in the motifs located in TMS1 and TMS2 result in decreased levels of Tim17 in complex with Tim23 (218) (Fig 3.1, Tim17, yellow). C^{118} and C^{120} have been shown to crosslink to TMS1 of Tim23 (219) (Fig 3.1, Tim17, pink). These structural elements and interactions give basis to the type of translocation pore that is formed between Tim23 and Tim17.

3.2.3 Coevolution Analysis Applications

Although these experimental studies above have give a number of structural insights into Tim23 and Tim17, there is no high-resolution structure of the channel

domain of these homologous proteins. Coevolution analysis has previously determined the structure of numerous membrane proteins (200). In this study, the sequence of the channel domain of homologous proteins, Tim23, Tim17, and Tim22 were used to determine a structural model of the channel domain.

Coevolution can also be used to determine interactions between two proteins (220). In this case, a paired protein alignment is created and the analysis not only identifies coevolving residue pairs within the proteins of interest but between the two proteins as well. This study utilizes coevolution analysis to also determine the coevolving pairs between Tim23 and Tim17, the main channel forming proteins of the TIM23 complex.

3.3 Materials and Methods

3.3.1 Alignment Generation, GREMLIN Server Analysis, and Rosetta Structural Modeling

A MSA of the Tim23 superfamily was created and included sequences for Tim23, Tim17, and Tim22. Sequences were collected from the NCBI Database (221) and UniProt Database (222). Alignments were supplemented by the GREMLIN webserver using HHblits (201,202) to add additional homologous sequences. Only hits with $1E^{-10}$ or below were considered and the alignment was processed so that all sequences had at least 75% coverage (no more than 25% gaps when aligned) and that all columns within the MSA that were more than 25% gaps were removed. The MSA was also filtered so that no two sequences had greater than 75% sequence identity. The GREMLIN webserver (192,196) was used as described in section 2.3.1. In collaboration

with Dr. Sergey Ovchinnikov, the coevolution analysis results were weighted and implemented as sigmoidal distance restraints for structural modeling with Rosetta (199).

3.3.2 ConSurf

ConSurf is a webserver that estimates the evolutionary conservation of an amino acid in a protein based on a MSA (223-227). The rate at which an amino acid is conserved depends on the structural and functional importance of the residue. The ConSurf webserver uses a Bayesian method that not only takes into consideration the variation of the MSA but also the phylogenetic relationship of the analyzed sequences (228). Each amino acid of the protein of interest is assigned a conservation score ranging from 1: variable, to 9: conserved, with 5 representing an average score.

3.3.3 Paired Protein Alignment Generation

Sequences designated as Tim23 and Tim17 of the same species were concatenated together. In this analysis, the channel domain of Tim23 and Tim17 were analyzed. In the case of *S. cerevisiae,* the Tim23-Tim17 concatenated sequence is 266 residues in length. The joint sequence consists of the C-terminal channel domain of Tim23 (residues 1-126 of the concatenated sequence) and the entire Tim17 sequence (residues 127-266 of the concatenated sequence). Once Tim23 and Tim17 sequences were concatenated, a MSA was produced, and submitted to the GREMLIN webserver for coevolution analysis.

3.3.4 RaptorX Contact

The RaptorX Contact webserver (197,198), as described in section 2.3.5, was used to analyze the Tim23 superfamily of proteins.

3.4 Results and Discussion

3.4.1 Coevolution Analysis of the Tim23 Superfamily

Coevolution analysis was performed and mapped onto the channel domain of Tim23. The alignment consisted of sequences from the Tim23 superfamily of proteins and contained 1,194 sequences resulting in a sequence/length ratio of approximately 9.5. The GREMLIN webserver predicted 179 coevolving pairs. The contact map of the coevolution analysis suggests antiparallel alpha-helical interactions as indicated by the perpendicular pattern off the diagonal (Fig 3.2). Analysis of the contact map and the coevolving predictions show strong coevolution between residues of H2 and H3 and H3 and H4. These antiparallel interactions are consistent with the predicted four transmembrane helical topology of Tim23.

3.4.1.2 Tim23 Superfamily Predicted Models with the GREMLIN webserver

Rosetta structural modeling of the Tim23 superfamily alignment of proteins generated over 20,000 models, with the top ten structures clustering into two groups. Cluster one contained six predicted models whereas cluster two contained two predicted models.

3.4.1.2.1 Model One

A representative model from cluster one is shown (Fig 3.3A, left). It contains five helical structural units, which agrees with the generally held belief that the channel domain of Tim23 is made up of four transmembrane helices and one α-helix located in loop 1 (L1) found in the matrix (Fig 3.3B). When compared to the predicted coevolving residues, the resulting Rc value is 0.631, suggesting not quite a native protein fit

between the predicted coevolving residues and model one but a better fit than comparison to a random protein.

The predicted coevolving pairs can be compared to the model in the form of a contact map. The blue dots represent the predicted coevolving pairs whereas the dark grey dots represent residues known to be in contact with each other in the model (Fig 3.4A, left). The highlighted ovals show the main structural interactions seen on this contact map. Here we illustrate how the predicted coevolving contacts and the known contacts of the structures are similar and different. For both the coevolving pairs and the contacts made in the model there is a strong pattern indicating H2:H3, H3:H4, and H1:L1 interactions. However, when looking at the known contacts made by model one there are interactions between H1:H3, H1:H4, and H2:H4 that are not represented by the coevolution analysis.

To better understand the coevolving interactions, the predicted pairs with a probability (of being within 5Å of each other) of greater than 0.9 were mapped onto model one (Fig 3.4B, left). This allowed us to visualize the distances and locations between the top predicted pairs on model one. The majority of the short distance pairs (<5Å apart) suggest the alpha-helical nature of the channel domain of Tim23, in that there are many pairs occurring in an i, i+3/i+4 pattern. Other short distance pairs appear between H2:H3 and H3:H4, which were noted on the contact map. There are also a number of intermediate distance (5-10Å) and long distance (>10Å) coevolving pairs, which may suggest structural dynamics, as we have shown for the AAC (Chapter 2). Specifically, the long distance connections highlight residues of L1 coevolving with

residues of the TMSs. We believe this suggests L1 inserts into the transmembrane domain.

3.4.1.2.2 Model Two

A representative structure from cluster two is shown (Fig 3.3A, right). Model two also contains five helical structural units, although this model suggests all five helical units are TMSs. What was identified in model one as the helical element of L1 in the matrix has now shifted into register with the other TMSs (Fig 3.3B). Previous research indicates this is not the topology of Tim23 due to both the N- and C- terminal ends of Tim23 being located in the IMS (211). We believe this model could be reiterating the suggested dynamic interaction between L1 and the TMSs as proposed by model one.

When compared to the predicted coevolving residues, the resulting Rc value is 0.642, again suggesting not quite a native protein fit between the predicted coevolving residues and structural model but a fit better than a random protein comparison. Differences between the known contacts of model two and the predicted coevolving pairs can be seen on the resulting contact map (Fig 3.4A, right). Similarities again highlight the anti-parallel interactions of the TMSs, specifically, H2:H3 and H3:H4. The most pronounced differences occur with the model suggesting interactions between L1 and H2, H3, and H4 that are not well represented by the coevolution analysis.

To better understand these differences and the potential structural dynamics, the coevolving pairs with a probability of greater than 0.9 were mapped onto model two (Fig 3.4B, right). These predicted contacts reiterate the incorrect prediction that Tim23 has five TMSs. The long and intermediate connections cross the center of the protein, with

residues on one side of the protein coevolving with residues on the complete other side of the protein. This model satisfies a number of coevolving pair predictions however, the number of long and intermediate coevolving pairs validate the topology of model one. This flipped topology of model two demonstrates the high coevolution between L1 and the transmembrane domain of Tim23 and supports the notion that L1 is structurally dynamic.

3.4.1.2.3 Model Comparison

Although both models have problematic aspects, they do generally agree with what we know about the structure of the channel domain of Tim23 and suggest the structurally dynamic nature of L1. Model one is problematic in that the register of H1 with the other TMSs is not energetically favorable if we consider how this protein would be positioned in the membrane. When considering model two, the problematic aspect is the topology. Structural modeling using coevolution analysis will attempt to satisfy the greatest number of predicted coevolving pairs. The two models attempt to satisfy the predicted structural dynamics of L1 in two different ways. In model one, this is done by shifting H1 out of register with the other TMSs and into register with the helix of L1. In model two this is done by inserting L1 into the transmembrane domain of Tim23 resulting in the inverted topology of H2, H3, and H4.

To gain insight into the structural features of this potentially dynamic 16-residue L1 helix identified by our coevolution analysis, we performed helical wheel analysis of this segment. Interestingly, this region defines a helix with a hydrophobic face flanked by basic residues (Fig. 3.5). Hence, this structural element shares physicochemical features with N-terminal presequences on polypeptides that traverse the Tim23

channel. Given the dynamic, channel-interactive nature of this segment identified by coevolution analysis, we propose that L1 may serve as a temporary presequence-like structure that occupies the channel lumen when the TIM23 channel is not engaged with substrate. Considering, the positively charged nature of the L1 helix, this interaction could be dependent on CL. The negatively charged head group of CL could attract the positively charged helix of L1, supporting these structural dynamics. This interaction should be investigated experimentally to determine the validity of the predicted structural dynamics and to determine the CL dependence of this effect.

3.4.1.3. Analysis of Highly Coevolving Residues

Coevolution analysis can be used to identify possible functionally important residues (200). We analyzed the results for the Tim23 superfamily to identify residues that might be of interest for future experimental studies. To determine the highly coevolving residues, the coevolution score was summed for each residue and designated as potentially functionally important if it appeared in the top 20% of residues (Fig 3.6).

We wanted to validate this approach by determining if any previously identified functionally important residues (Fig 3.1) were also identified by coevolution analysis alone. Residues Y105 to F114, of H1, can be crosslinked to residues of Tim17 (219), this coevolution analysis identified L106 and G108 as functionally important. A lethal phenotype results from the mutation G149L (213), and this glycine residue appeared in this analysis. Finally, G145 has been determined to stabilize the interaction between Tim23 and Tim17 (213); this glycine residue was also identified in this analysis. These

examples lend confidence to this approach for identifying highly coevolving residues as functionally important.

Analysis of how these residues map onto the various structural elements shows the greatest number of identified residues map to L1, underscoring the potential functional significance of interactions of this region with the channel domain. Interestingly, the high coevolution score of this structural element is also consistent with those of the matrix-facing helices of AAC; the latter may be related to the similar topological locations of these structural elements relative to the membrane surface. It is important to consider these residues are mapped onto the Tim23 sequence, however this analysis was done using all members of the Tim23 superfamily (Tim23, Tim17 and Tim22). These highly coevolving residues could be important for the function of any of the three homologous proteins.

3.4.1.4 Conservation Analysis of the Tim23 Superfamily

The ConSurf webserver was used to complete a conservation analysis of the Tim23 superfamily. Graphing the conservation analysis along each TMS shows clearly how one face of each helix is more conserved whereas another is more variable. The clear oscillation of the graph indicates that within all the TMS there is an i, i+4 pattern for highly conserved residues (Figure 3.7A). This pattern identifies a polar inward facing side of H2 in model one that previous research has identified as part of the protein translocation pore (214) (Fig 3.7B, left). The highly conserved nature of the other TMSs suggests their importance in translocation pore formation. The identified variable face (Fig 3.7B, right) could interact with the hydrophobic membrane and therefore not be subject to strong evolutionary constraints. However, it could also be the case that a

more variable face is necessary to facilitate the number of different interactions with other subunits of the TIM23 complex necessary for functional protein import or integration.

3.4.1.5 Structure Prediction with the RaptorX Contact Webserver

As a second method for both coevolution analysis and structure prediction, the RaptorX Contact webserver was used to predict the structure of Tim23. The input sequence for the RaptorX Contact prediction was the channel domain of Tim23 from *S. cerevisiae*. RaptorX Contact predicted five possible models, all of which were very similar in overall structure to the GREMLIN model one structure (Fig. 3.8A). This RaptorX Contact webserver predicted Tim23 has four TMSs and an α-helix in L1 in the matrix. Looking at the top predicted contacts for the RaptorX Contact model of Tim23 indicates the anti-parallel nature of the channel domain of Tim23 with the top coevolving pairs occurring between neighboring TMSs (Fig 3.8B). This model also suggests the interaction of H1 with L1 as a number of top coevolving pairs occur here as well. The agreement of this RaptorX Contact model with the GREMLIN model one suggests this could be a useful structure to understand the channel domain of Tim23 gives us confidence of the structure and dynamic nature of the α-helix in L1.

3.4.2 Coevolution Analysis of the Tim23 and Tim17 Interaction

The resulting Tim23-Tim17 alignment consisted of 421 concatenated sequences. With an alignment-filtered length of 248 the resulting sequence/length ratio was 1.7. The coevolution analysis resulted in 372 predicted coevolving pairs. Resulting pairs are shown in a contact map (Fig 3.9A) where the top left and bottom right quadrants

indicate coevolving pairs within Tim23 and Tim17 respectively, whereas the bottom left and top right quadrants indicate coevolving pairs between the two proteins. Of note are two high scoring coevolving pairs between Tim23 and Tim17 (Fig 3.9A, circled pairs). What is most promising is these predicted interactions are topologically possible. The pairs of note are K190 (Tim23) with K55 (Tim17) and S192 (Tim23) with K55 (Tim17). K190 and S192 are both located in L3 in the matrix and K55 of Tim17 is located in L1, also in the matrix (Fig 3.9B). RaptorX Complex Contact webserver was used as a second means to validate the coevolution results and predicted these same two interactions. Perhaps most striking is that no transmembrane packing interactions are identified by this coevolution analysis. Tim23 and Tim17 are believed to constitute part of the protein conductance channel. It has therefore been speculated, although never shown experimentally, that these two homologous proteins may assemble through helical packing interactions into a common channel. This analysis suggests that Tim23 and Tim17, while both having strong patterns of intramolecular coevolution interactions, have very little intermolecular coevolution. Where detectable coevolution does exist in soluble loops between TMSs. Hence, this analysis would suggest that the assembly of Tim23 and Tim17 is not driven by robust transmembrane helical interactions

3.4.3 Conservation Analysis of the N-Terminal Domain of Tim23

The N-terminal domain of Tim23 (residues 1-96 in *S. cerevisiae*) is not conserved and does not align well in a MSA, making it unsuitable for coevolution analysis. However, it has been shown that a series of residues, W3-L4-F5, are important for the interaction between Tim23 and the OM and that this interaction is strengthened by the presence of CL (102). This series of residues is known as the hydrophobic hook. With

the extensive database of Tim23 sequences collected we wanted to evaluate the conserved nature of these residues and if there were any other defining features of the N-terminal domain.

Tim23 sequences from fungal species were collected and resulted in a database of 357 sequences. To understand if the hydrophobic hook is highly conserved, we determined how many sequences contained this motif. Surprisingly, only 12% of sequences contained the exact W-L-F motif (Fig 3.10A, Hydrophobic Hook). This led us to hypothesize that it is not the exact sequence of the hydrophobic hook that is important for OM interaction. Rather, it is possible that the presence of specific amino acids within the extreme amino terminus may be the defining characteristic of this region. To this end, the database was analyzed for the presence of certain amino acids within the first ten residues of the N-terminal domain of Tim23. What was found illuminated that there is an undeniable prevalence of certain amino acids within these first ten residues. The most defining characteristic of this region was the presence of an aromatic residue (86% of all cases), which was most commonly tryptophan but could also be phenylalanine (Fig. 3.10A).

The sequence of the first ten residues of the N-terminal domain of Tim23, M-S-W-L-F-G-D-K-T-P (*S. cerevisiae*), indicates that following the hydrophobic hook residues, there is a glycine residue followed by an aspartic acid residue and then a lysine residue. The same analysis was performed to see if the prevalence of a glycine residue, negatively charged residue (D/E), or positively charged residue (K/R) was conserved. Upon analysis, the glycine residue is present in 78% of sequences, the negatively charged residue, 62% of sequences, and the positively charged residue 63%

of sequences (Fig 3.10A). Finally, when looking at the N-terminal ten residues of the different Fungi species there was an obvious indication that the majority of sequences had a serine residue in the second position. Quantifying this observation showed that 67% of sequences have a serine residue at the second position (Fig 3.10B) indicating the highly conserved nature of this residue and position. Together, this analysis indicates it may not strictly be the hydrophobic hook residues that are necessary for OM interaction. Instead, we believe it is the presence of an aromatic residue, a glycine residue, a positively charged residue, and a negatively charged residue along with a serine residue in the second position that are important for Tim23 interaction with the OM.

3.5 Conclusions

Membrane proteins have been notoriously hard to study by traditional structure determination methods, however with the advent of high throughput sequencing technologies a greater number of proteins have become amenable to structural bioinformatics-based approaches such as coevolution analysis. Not only have more proteins been sequenced, providing a richer database to work with but there have been vast improvements in these methodologies as well. Additionally, coevolution now has the ability to utilize the contact predictions made to inform *ab initio* structural modeling. We have used coevolution analysis to determine novel insights into the channel domain of the Tim23 superfamily. We have utilized these coevolution results to perform *ab initio* structural modeling to determine the first model of the channel domain of Tim23. In future studies this model could be used to supplement traditional structure

determination methods that would aid in the determination of the first high-resolution structure of the Tim23, Tim17, or Tim22 channel domain (220).

Analysis of the distances of the coevolving residue pairs illustrates the potential structural dynamics of the matrix helix of L1. Together with the hydrophobic and basic residues present in the L1 helix we propose this structural element acts as a presequence-like structure to close the translocation channel when TIM23 is not actively engaging precursor protein. This analysis also identified potentially functionally important residues based on their highly coevolving nature. Some of these identified residues have been the targets of past experimental studies and have shown their importance for either the interaction between Tim23 and Tim17 or resulted in a lethal phenotype lending credence to this analysis. The remainder of these identified residues should be investigated to determine if there is a functional impact of mutation both in the context of protein function and TIM23 complex assembly. Included in this analysis should be the coevolving residues identified between Tim23 and Tim17.

Moving forward the working database of all sequences collected for Tim23, Tim17 and the majority of the other subunits of the TIM23 complex should continue to be curated (Appendix B). As more sequences become available, coevolution can be used in multiple ways to understand the structural elements and dynamics of subunits and interactions of the subunits of the TIM23 complex and can be used to inform experimental analysis and structure determination.

Lastly, the N-terminal domain of Tim23 is highly variable and therefore not suitable for coevolution analysis. To circumvent this and investigate the interaction of the Tim23 with the OM we analyzed the first ten residues of Tim23 to determine if there

was an evolutionary basis behind the concept of the “hydrophobic hook”. This unique analysis highlighted that the presence of an aromatic residue, a positively charge residue, a negatively charged residue, a glycine residue, and a serine in position two are conserved in our extensive database as opposed to the “hydrophobic hook” motif. Currently, The results from this analysis are being used by M. Skoryk of the Alder laboratory to inform fluorescent studies investigating how the mutation of these residues affects protein association with liposomes of various compositions.

3.6 Figures and Table

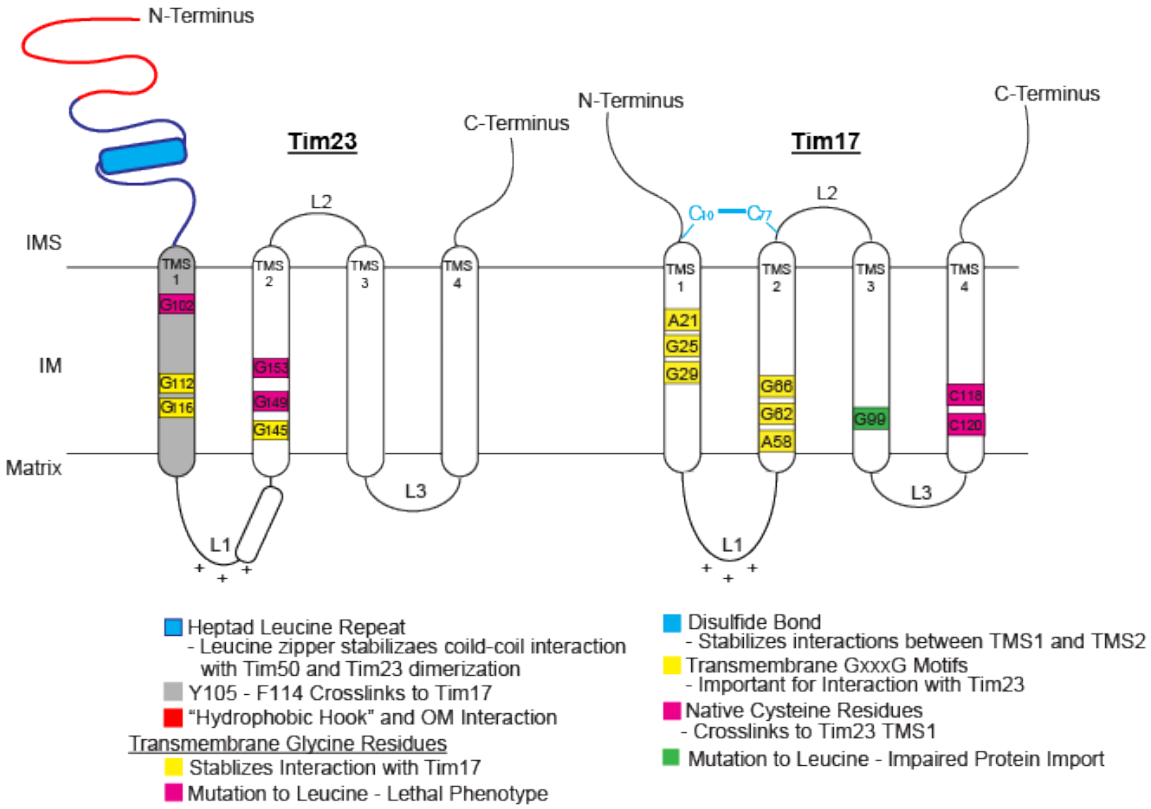


Figure 3.1. Schematic of Tim23 and Tim17 protein structural elements.

Tim23 and Tim17 are homologous proteins. *Left*, Tim23 consists of a soluble IMS N-terminal domain (residues 1-96, *S. cerevisiae*) and IM bound C-terminal channel domain (residues 97-222, *S. cerevisiae*). *Right*, Tim17 lacks the extensive N-terminal domain found in Tim23 but has a similar C-terminal channel domain. Residues and structural motifs that have been experimentally shown to be critical for the functional interactions of Tim23 and Tim17 are color coded and detailed on the key below each protein.

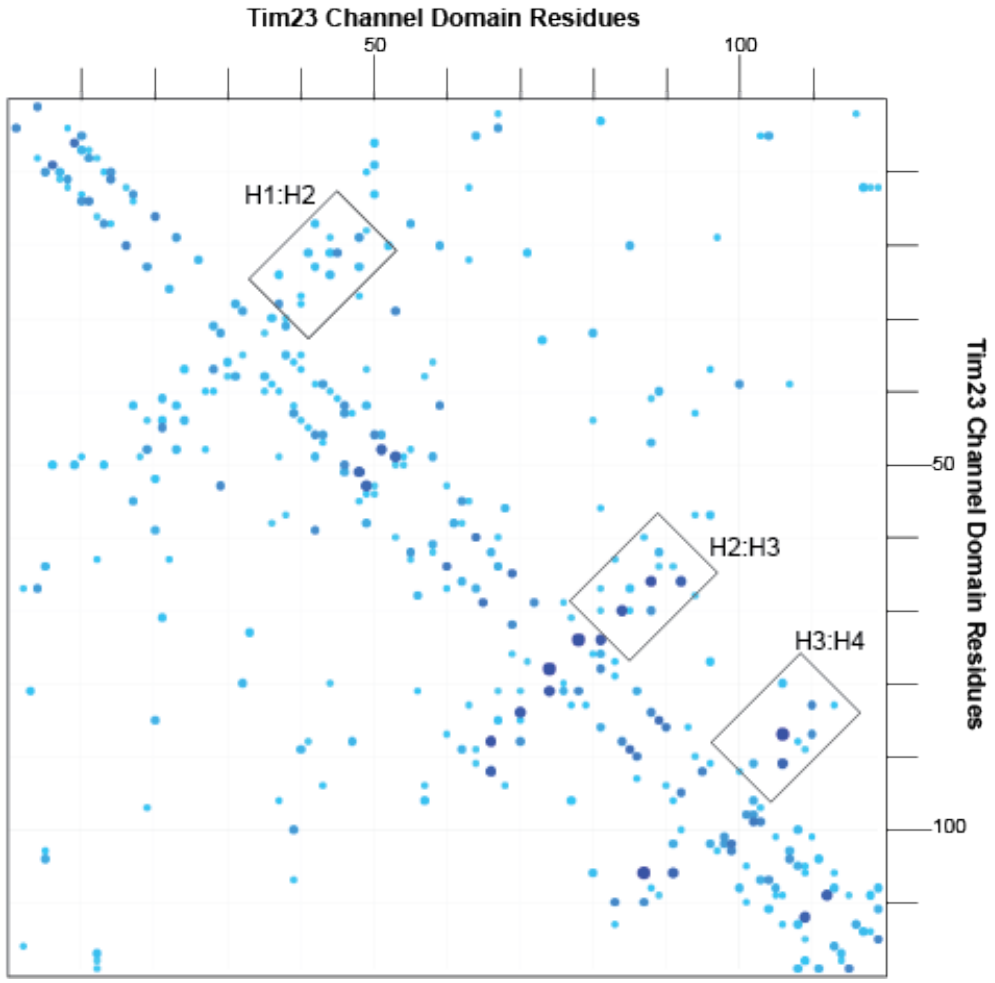


Figure 3.2. Tim23 superfamily channel domain contact map analysis using the GREMLIN webserver. The channel domains of Tim23, Tim17 and Tim22 were aligned and submitted for analysis through the GREMLIN webserver. Results were mapped onto the sequence of the Tim23 channel domain from *S. cerevisiae* and represented on a contact map with the residue of the channel domain of Tim23 (1-126) on each axis. Each blue dot indicates a pair of residues predicted to be coevolving, the darker and larger the dot, the higher the predicted coevolving strength. Rectangles highlight the H2:H3 and H3:H4 interactions.

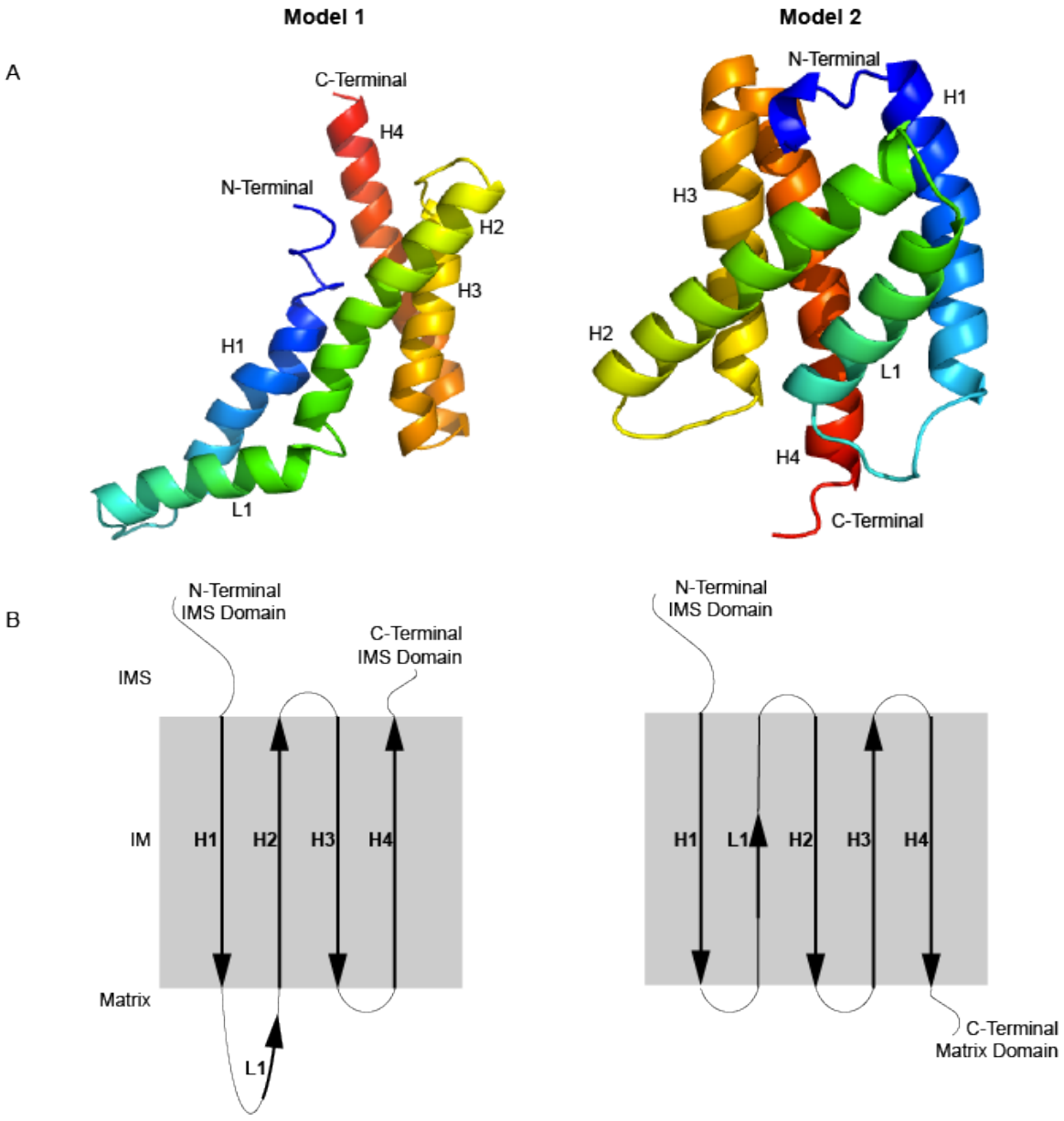


Figure 3.3. Tim23 superfamily predicted models using the GREMLIN webserver. (A) Structure of predicted models. (B) Topology diagram of the predicted models. *Left*, model one. *Right*, model two.

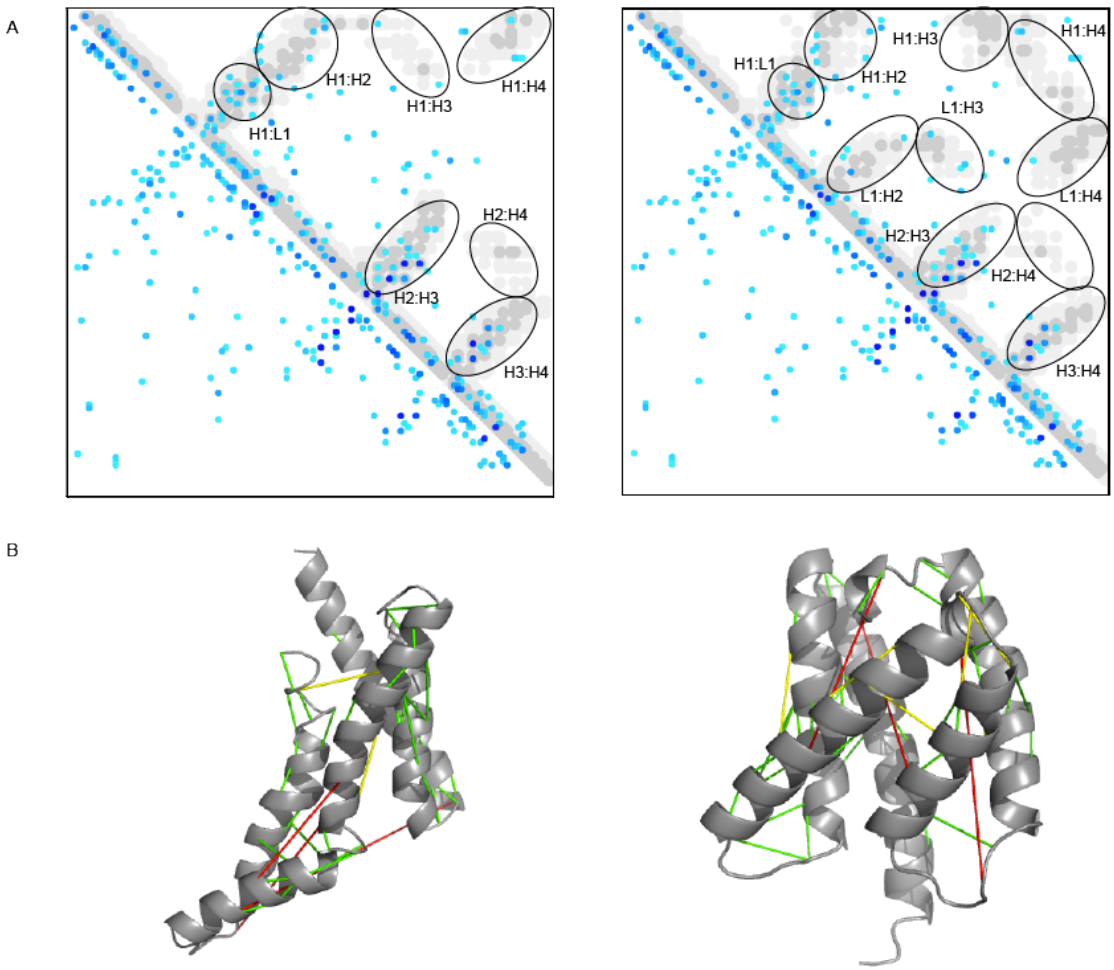


Figure 3.4. Comparison of predicted coevolving pairs to Tim23 structural models. (A) Contact maps with Tim23 residues on both axes. Each blue dot indicates a pair of residues predicted to be coevolving, the darker and larger the dot, the higher the predicted coevolving strength. Dark grey dots on the upper half of the contact map indicated residue pairs known to be within a distance of 5Å with light grey dots indicating residue pairs known to be 5Å-10Å apart. Ovals indicate interactions between the major structural units of the predicted models. (B) Top coevolving predicted hits were mapped onto the structure. Distances between residues are depicted as color-coded lines (green, <5Å apart; yellow, 5-10Å apart; red, >10Å apart). *Left*, model one. *Right*, model two.

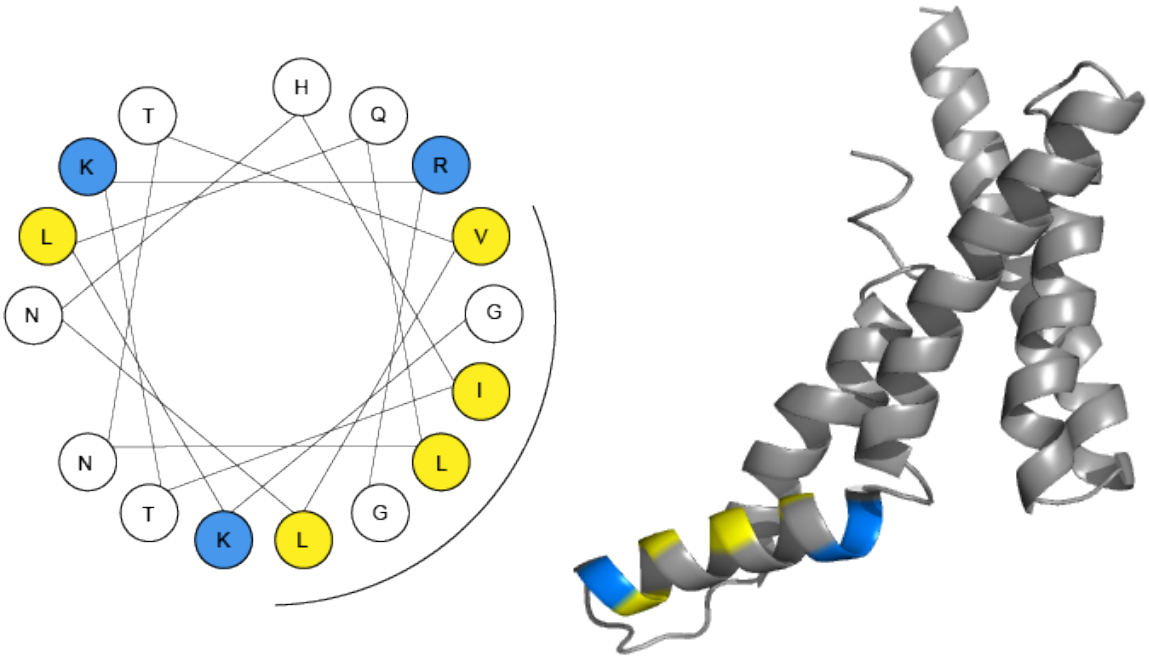
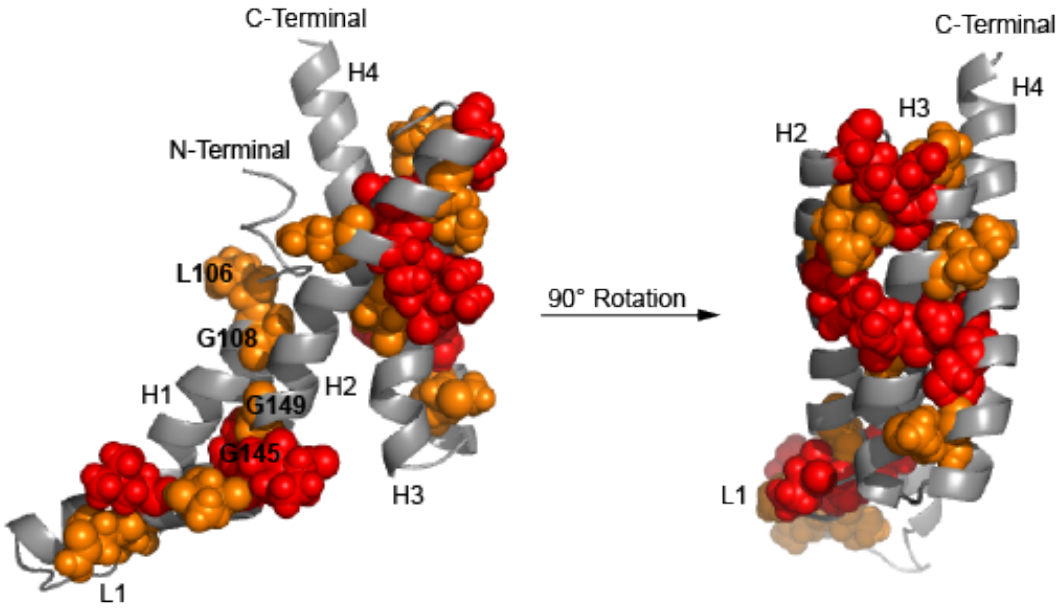


Figure 3.5. Property analysis of the L1 helix. Helical wheel analysis of L1 shows properties similar to an amino-terminal presequence with positively charged basic residues and a hydrophobic face. *Left*, helical wheel analysis. *Right*, residues of interest shown on structure. Coloring: yellow, hydrophobic; blue, basic.

A



B

Structural Region	Top 10%	Top 11-20%
H1	-	L106, G108
L1	L138, R144, G145, P146	N135, T136, T142
H2	I162, N163, R170	G149, N160, I166
H3	G177, A183, L184	D174, I179, T185
H4	S202, V205	M198

Figure 3.6. Highly coevolving residues. (A) The summed coevolution score of individual residues reflect potential functionally important residues. Top 10%, red spheres; top 11- 20%, orange spheres. (B) Table of all top 20% potential functionally important residues separated by predicted structural region of model one of the channel domain of Tim23.

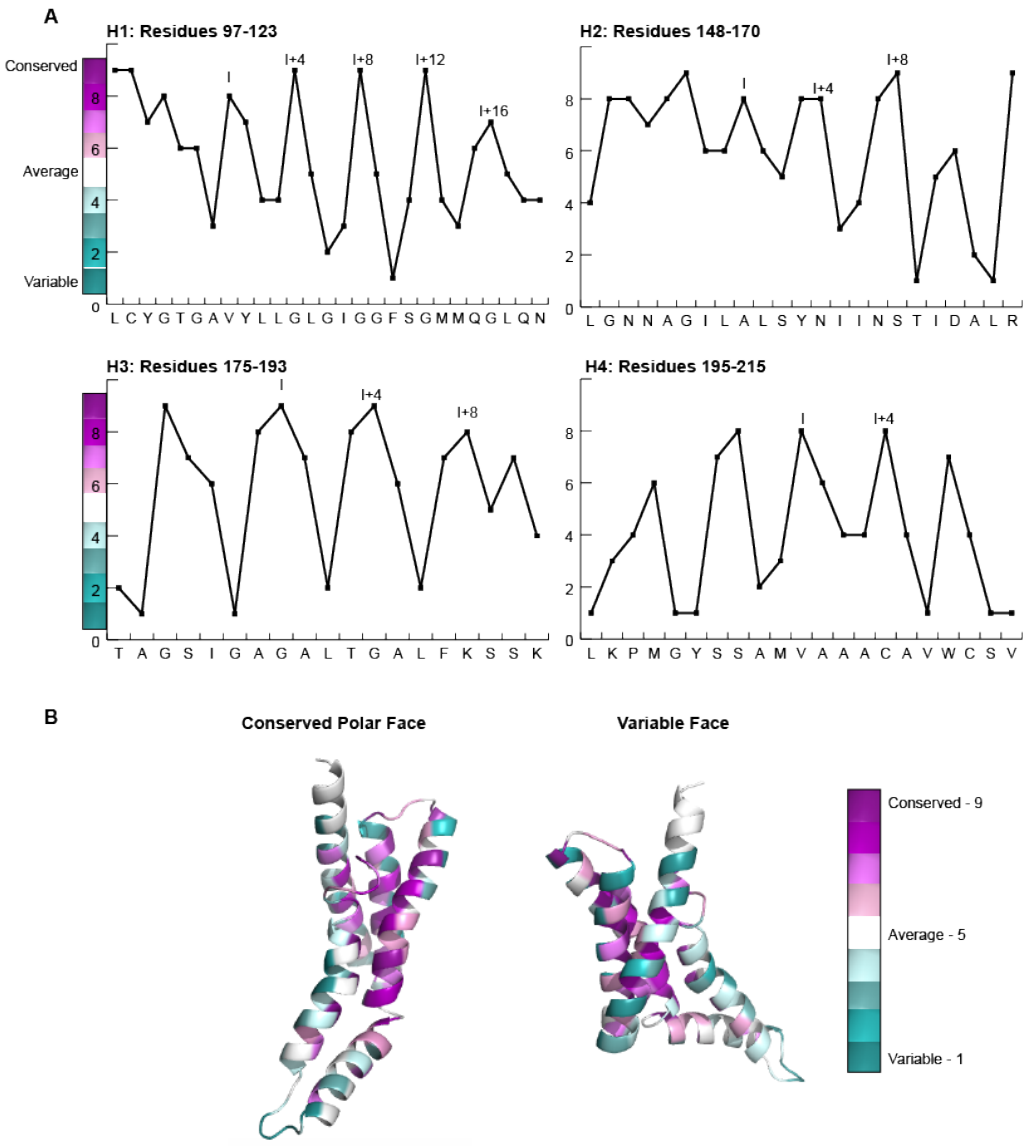


Figure 3.7. Conservation analysis of Tim23 with the ConSurf webserver. (A) Residues of the major structural elements are plotted on the x-axis and ConSurf score is plotted on the y-axis to show the periodicity of highly conserved residues. (B) Structures have been colored

and mapped onto model one according to ConSurf score to highlight the conserved polar face (left) and variable face (right).

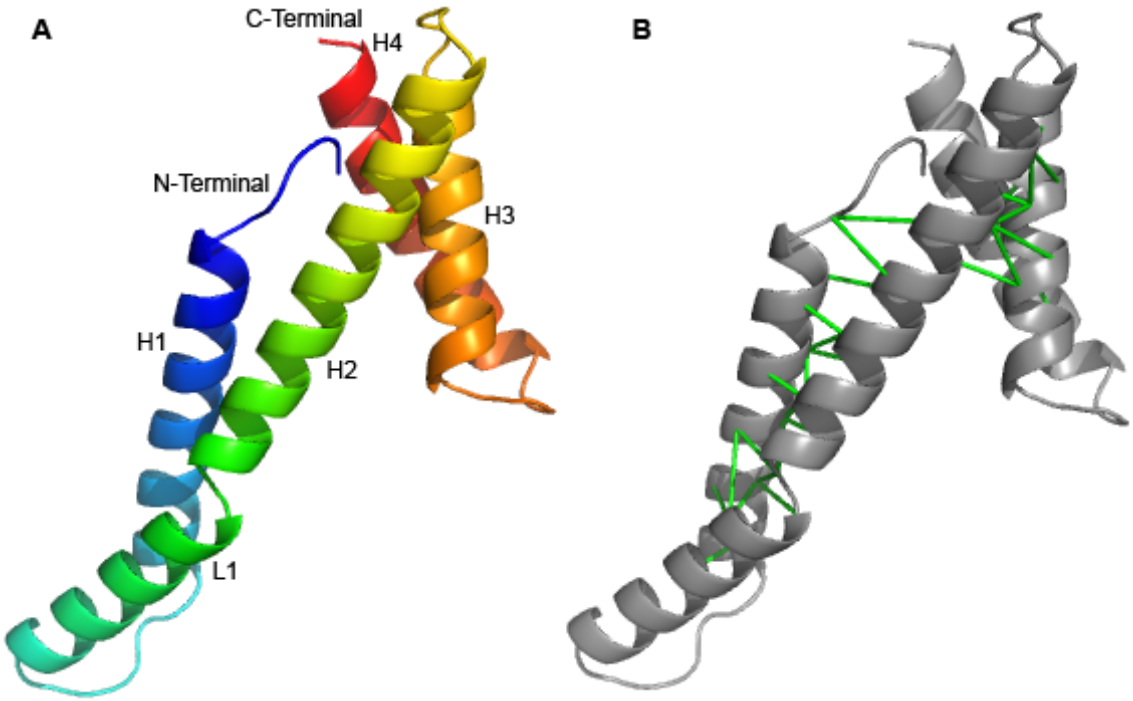


Figure 3.8. Tim23 model predicted by the RaptorX Contact webserver.

(A) Predicted structure of the channel domain of Tim23 by RaptorX Contact. (B) Top scoring predicted coevolving pairs (>0.9 probability score) were mapped onto the predicted structure. Indicated connected residues are predicted to have their C-beta atoms within 8Å of each other.

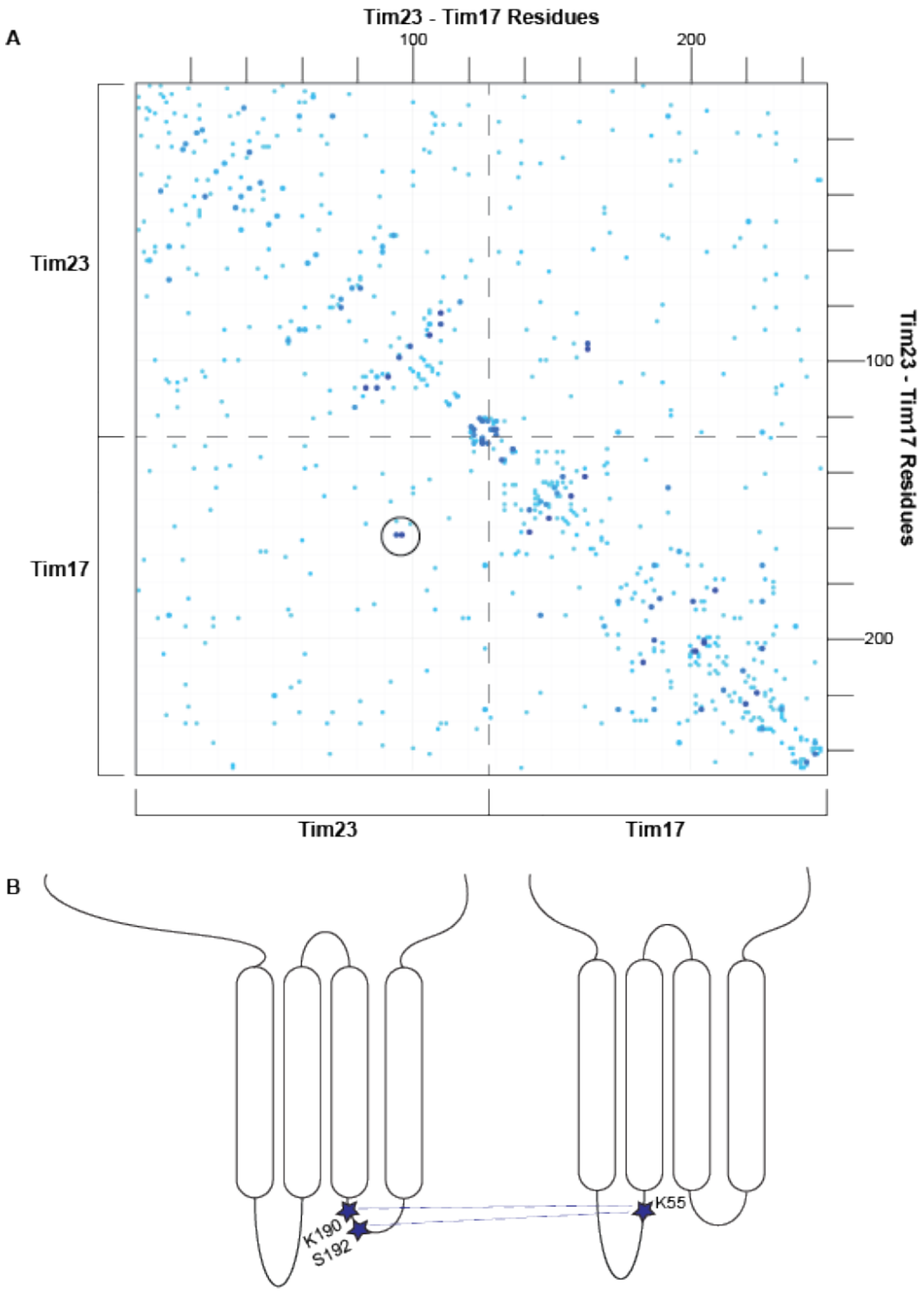


Figure 3.9. Tim23-Tim17 interaction analysis by the GREMLIN webserver. (A) Contact map of the Tim23-Tim17 predicted coevolving pairs. The sequence of the channel domain of Tim23 (residues 97-222) from

S. cerevisiae was concatenated with the sequence of Tim17 from *S. cerevisiae*. The residues of this concatenated sequence are plotted on the x- and y-axes. Each blue dot indicates a pair of residues predicted to be coevolving, the darker and larger the dot, the higher the predicted coevolving strength. Circled coevolving pairs indicated two high probability coevolving residue pairs between Tim23 and Tim17. (B) Schematic representation of the two predicted coevolving pairs between the channel domain of Tim23 and Tim17.

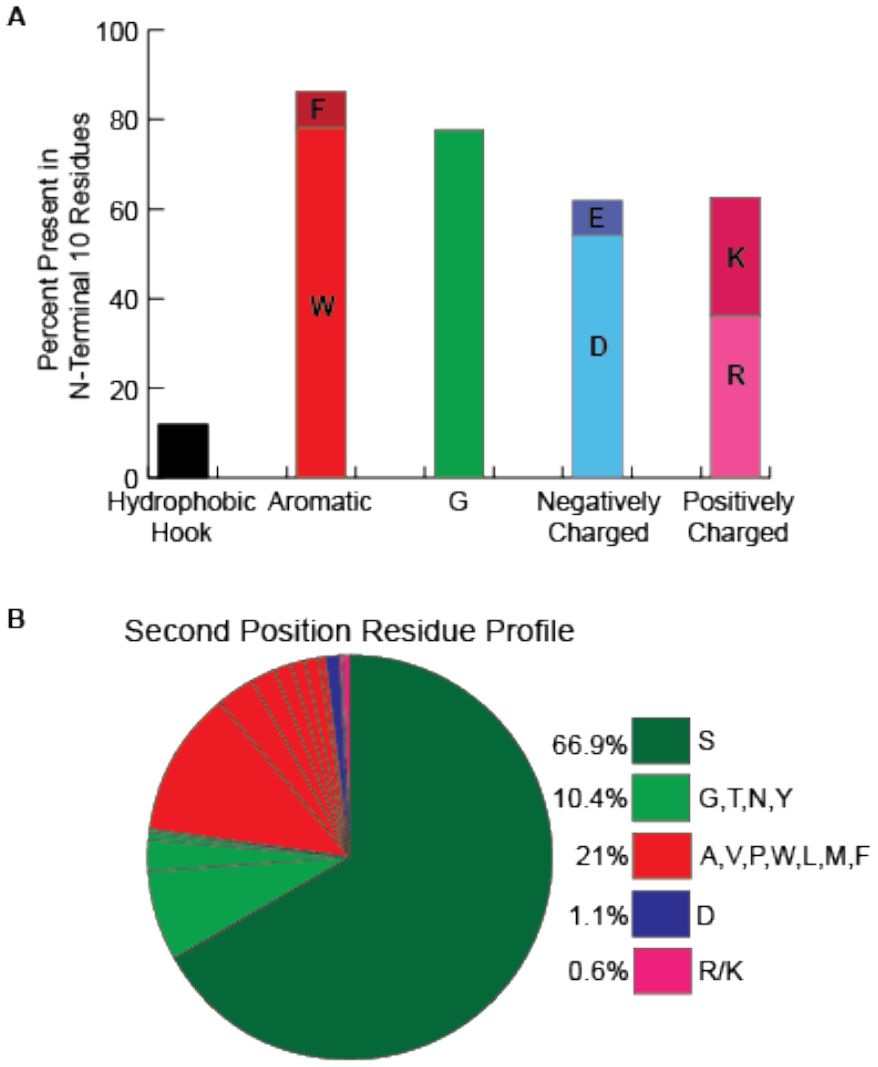


Figure 3.10. Conservation analysis of the N-terminal ten residues of Tim23. (A) The presence of certain amino acids within the first ten residues of Tim23 was determined. Results are shown in terms of percent of all sequence analyzed. (B) To understand the interaction between Tim23 and the OM, the amino acid profile of the second position of Tim23 was analyzed. Amino acids are color coded by type and presented in order in the pie chart.

Chapter 4

Understanding the Effects of Cardiolipin on the Distribution of TIM23 Assembly States

4.1 Abstract

The modular TIM23 complex contains 11 identified subunits that can assemble into distinct functional states that facilitate the membrane integration of IM proteins and the translocation of soluble proteins into the matrix. Due to this modular nature subunit assembly must be dynamic to support multiple functional states. Here we investigate the role that membrane lipids, specifically cardiolipin, play in modulating the TIM23 assembly states. Cardiolipin is the signature phospholipid of the mitochondria and is known to support proper protein activity and complex assembly of a multitude of mitochondrial proteins. Here we provided a comprehensive analysis of the distribution of functional TIM23 assembly states in mitochondria from yeast with altered cardiolipin compositions including complete loss of cardiolipin (Δ*crd1*) and accumulation of both unremodeled cardiolipin (Δ*cld1*) and monolysocardiolipin (Δ*taz1*). We find there is a destabilization of the $TIM23^{SORT}$ complex that results from the loss of cardiolipin or accumulation of monolysocardiolipin. We find that these alterations are not due to the decrease in membrane potential seen with Δ*taz1* and Δ*crd1.* Finally, we show preliminary results that suggest the destabilization of the $TIM23^{SORT}$ complex is reversible with the addition of supplemental CL in the form of CL containing NDs. We propose CL containing NDs could be implemented as a mitochondria specific targeted lipid delivery system.

4.2 Introduction

4.2.1 The TIM23 Complex Assembly States

The subunits of the TIM23 complex must all communicate effectively to ensure proper protein import. The various TIM23 assembly states have been experimentally

identified, primarily by native gel electrophoresis (37,56,93,229). Blue-native PAGE (BN-PAGE) is a technique that resolves protein complexes with their tertiary and quaternary structures intact. The use of a native environment and non-denaturing detergents allows visualization of oligomeric states, protein-protein interactions, and supercomplex formation (230). Previous research has identified three main assembly states of the TIM23 complex that can be visualized through BN-PAGE. The first complex referred to as the $TIM23^{CORE}$ complex, consists of Tim23 and Tim17 (72). Opposing research claims this complex may include Tim50 as well (229). The two higher molecular weight assembly states are referred to as $TIM23^{SORT}$ as they are known to contain Tim23, Tim17, and Tim21. Previous research has also claimed one of the $Tim23^{SORT}$ assembly states additionally contains Tim50 (37,56,72,229). Together these multiple assembly states demonstrate the modular nature of the TIM23 complex. Additionally, it has been shown these TIM23 complex assembly states are stabilized by CL (37,95,98,229) however, a number of questions remain in the literature about the extent of this effect.

4.2.2 Nanodiscs: A Model Membrane System

Nanodiscs (ND) are nanoscale lipid bilayers surrounded by membrane scaffolding protein (MSP), which renders them soluble and amenable to solution-based biophysical analyses that more accurately mimic the native membrane environment. Importantly NDs are stable and monodispersed in solution (231). Moreover, because they do not form closed structures like other model membranes such as liposomes, they allow experimental access to both sides of the bilayer (232). Traditionally, membrane proteins have been reconstituted into NDs and studied in a number of ways, some of

these membrane proteins include, bateriorhodopsin, proteorhodopsin, SecYEG, Tim50, Tim23, and complex II (98,215,233-237).

MSP was developed by the Sligar lab at the University of Illinois at Urbana-Champaign (232) and consists of a hexa-histidine tag, cleavage site, and variable lengths of amphipathic α-helices to produce NDs of a desired size (Fig 4.1). By engineering constructs consisting of one (MSP1E1), two (MSP1E2), or three (MSP1E3) copies of the original MSP, different size NDs can be produced. During development it was discovered 22 N-terminal amino acids were not essential for nanodisc formation and did not affect the resulting ND size, the deletion of these residues is denoted with the addition of “D1”. MSP1E3D1, the form used in this study, contains three copies of the original MSP protein with the N-terminal 22 residues deleted, and results in a ND with a diameter of about 12nm (238). NDs are formed from a reaction mixture composed of MSP and lipid films, of a desired composition, and hydrated in detergent-containing buffers. NDs self-assemble upon removal of the detergent and then are purified by affinity chromatography (238).

4.2.3 The Effect of Supplemental CL on Mitochondria

Previous research has shown that addition of CL, in lipid micelles or assembled in NDs, to isolated mitochondria or cells results in the incorporation of the supplemental CL into mitochondrial membranes (239,240). Specifically, the addition of CL micelles has been shown to stabilize respiratory supercomplexes in Δ*crd1* yeast mitochondria (240). In an additional study, CL containing NDs were added to myeloid progenitor cells and shown to localize to the mitochondria and prevent the apoptotic response seen in TAZ knockdown cells (239). These results suggested that the addition of supplemental

CL to mitochondria or cells can reverse the effects of the lack of CL. Due to the modular nature of the TIM23 complex this study aimed to provide a comprehensive analysis of the effect of CL on the TIM23 complex and then use this experimental readout for understanding the effect of supplemental CL on the TIM23 assembly states and mitochondria in general.

4.3 Materials and Methods

4.3.1 Mitochondria Isolation

Mitochondria were isolated from yeast strain YPH499. Regardless of CL background all were isolated following the same procedure, as previously described (241). All *S. cerevisiae* cultures were grown in a lactate rich media at 30°C to an OD_{600} between 0.8 and 1. Mitochondria were quantified by A_{280} and stored at -80°C until needed, but no longer than four months.

4.3.2 Recrystallization of Digitonin

Commercially available digitonin (Sigma # D5628) was recrystallized to ensure purity. Digitonin was dissolved in 100% ethanol in a glass beaker and heated to 120°C, and boiled for 15 minutes to ensure all digitonin was dissolved in solution. The solution was then transferred equally to Nalgene Oak Ridge High-Speed Polycarbonate Centrifuge Tubes and cooled sequentially to allow for digitonin precipitation, at RT, 4°C, and -20°C, for 20, 25, and 30 minutes respectively. Digitonin was collected by centrifugation (15 minutes at 4°C and 7,000xg). The supernatant was discarded and digitonin was again dissolved in 100%, heated to 120°C, and heated until the digitonin was full dissolved. The solution was then again cooled sequentially as noted above but left overnight at -20°C to ensure optimal precipitation. Digitonin was collected by

centrifugation as noted above and dried using an Eppendorf Vagufuge Plus for six hours on the D-AL setting.

4.3.3 Mitochondria Solubilization and Preparation for BN-PAGE

Mitochondria were prepared for BN-PAGE as previously described (230). Briefly, mitochondria were thawed on ice for 30 minutes before centrifugation (five minutes at 4°C and 16,000xg) and were then resuspended in resuspension buffer (20mM HEPES pH 7.5, 80mM KCl, 3mM $MgCl_2$, 250mM sucrose, 2mM KPO_4 pH 7.4). Mitochondria were quantified with an A_{280} reading and adjusted to the desired concentration (4mg/mL). Mitochondria were aliquoted into the desired reaction volume and treated with 2mM NADH. In experiments requiring membrane potential disruption, carbonyl cyanide 3-chlorophenylhydrazone (CCCP) was added at various concentrations. Mitochondria were collected by centrifugation (five minutes at 4°C and 16,000xg) and resuspended in solubilization buffer (1% digitonin, 20mM Tris HCl pH 7.5, 60mM NaCl, 0.1mM EDTA pH 7.5, 10% glycerol, 1.6mM phenylmethylsulfonyl fluoride). Mitochondria were solubilized for 30 minutes at 4°C on a rotisserie. Samples were centrifuged (thirty minutes at 4°C and 16,000xg) and solubilized complexes were used for analysis or kept at 4°C overnight if necessary.

4.3.4 BN-PAGE Analysis

4-15% Mini Protean TGX Stain-Free precast protein gels (Bio-Rad #4568084) were used for BN-PAGE analysis with the Mini-PROTEAN Tetra Vertical electrophoresis cell. Solubilized mitochondria were combined with sample buffer (10mM Bis-Tris/HCl pH 7.0, 50mM 6-aminocaproic acid, 0.5% Coomassie brilliant blue (CBB) G-250, 5% glycerol) and loaded on the gel. NativeMark Unstained Protein Standard

(ThermoFisher #LC0725) was used as a molecular weight marker. Cathode buffer+CBB (50mM tricine, 15mM Bis-Tris/HCl pH 7.0, 0.02% CBB G-250) and anode buffer (50mM Bis-Tris/HCl pH 7.0) were loaded and the gel was run at 100V until the samples were well within the gel and then run at 20mA for the remainder of the electrophoresis step. When the dye front reached about two thirds of the way down the gel, the cathode buffer was replaced with cathode buffer lacking CBB (50mM tricine, 15mM Bis-Tris/HCl pH 7.0).

Western blotting was performed using the GenScript eBlot L1 wet protein transfer system. PVDF membranes were activated in methanol for five minutes and equilibrated in eBlot L1 PVDF Membrane Equilibration Buffer for ten minutes. One five-minute transfer cycle and one thirty-second cooling cycle was used to transfer proteins. After transfer, membranes were soaked in methanol for thirty minutes, with methanol changed every ten minutes. Membranes were then treated with primary antibody, αTim23, αTim17, αTim50 (provided by Dr. Dejana Mokranjac), or αTim21 (provided by Dr. Toshiya Endo) and then treated with a Cy5-conjugated secondary antibody. Blots were imaged using a Bio-Rad Imager under the Cy5 setting (635nm laser, 695nm emission filter, low sample intensity). Quantification of western blot images was done with Bio-Rad Image Lab (Version 3.0). To determine the proportion of the protein of interest in each of the TIM23 assembly states, one assembly state was set as the reference band; the relative intensity of the other assembly states was then compared to this. For Tim23 and Tim17, the reference band was the $Tim23^{CORE}$ assembly state. For Tim21 and Tim50, the reference band was the $TIM23^{SORT*}$ assembly state. These values were added together and the relative proportions of each assembly state were

determined as a percentage of the total. Statistical analyses were done with GraphPad Prism (Version 7.04).

4.3.5 2D BN/SDS-PAGE Analysis

For 2D BN/SDS-PAGE analysis the completed BN-PAGE gel was treated with a 1% SDS, 1% β-mercaptoethanol solution for 30 minutes at 50°C, with agitation every ten minutes. The gel was washed two times with 1X SDS running buffer for five minutes each. After treatment, each lane was cut and loaded onto a 15% SDS-PAGE gel that was poured using the Mini-PROTEAN comb Prep+1 (Bio-Rad #1653367). The gel slice was covered with stacking gel and allowed to polymerize. The gel was run at 120V for about two hours. 2D BN/SDS-PAGE gels were analyzed by western blot analysis as described above.

4.3.6 ND Assembly Reaction

MSP1E3D1 was purified as previously described and used for ND assembly reaction (232,238). Briefly, MSP was added to lipid films hydrated in a cholate-containing buffer (59mM). ND assembly was initiated upon removal of the cholate through the addition of Bio-Beads (Bio-Rad #1523920); the reaction was incubated with the Bio-Beads for two hours at room temperature. The reaction mixture was separated from the Bio-Beads and then incubated with Ni-NTA agarose resin for 45 minutes at room temperature and assembled NDs were purified by utilizing the hexa-histidine tag located on MSP through the addition of increasing concentrations of imidazole. Reaction buffer was exchanged to remove excess imidazole and the ND sample was concentrated using an Amicon 3KDa molecular weight cutoff filter (Millipore #UFC8003).

The ND lipid compositions used for these studies were POPC only and 50mol% POPC/50mol% TOCL.

4.3.7 CL Supplement Sample Preparation

Mitochondria were quantified, aliquoted, and treated with NADH as noted in section 4.2.3. Mitochondria were collected by centrifugation (five minutes at 4°C and 16,000xg) and resuspended in solubilization buffer lacking digitonin. ND were added at varying concentrations to mitochondria and were incubated on a rotisserie at 4°C for thirty minutes. Digitonin was added at a final concentration of 1% and samples were incubated again on a rotisserie at 4°C for thirty minutes. Solubilized complexes were collected by centrifugation (thirty minutes at 4°C and 16,000xg) and were kept at 4°C overnight if necessary, and subsequently used for BN-PAGE analysis.

4.4 Results and Discussion

4.4.1 CL Stabilizes Particular TIM23 Assembly States

Following the electrophoretic separation of digitonin-solubilized mitochondria by BN-PAGE, the TIM23 complex was shown to migrate as three distinct assemblies, $TIM23^{CORE}$ containing Tim23 and Tim17, and two higher molecular weight assemblies of $TIM23^{SORT}$ that we denote as $TIM23^{SORT}$ and $TIM23^{SORT*}$. These assemblies range in size from 90kDa to approximately 300kDa. To obtain a comprehensive analysis of the effect of CL on the stabilization of the TIM23 complex we evaluated the TIM23 complex composition in mitochondria from wild type (WT), Δ*taz1*, Δ*crd1*, and Δ*cld1* yeast strains (Fig 1.4).

The most dramatic effect on TIM23 assembly caused by defects in CL biogenesis was observed for the $TIM23^{SORT*}$ complex (Fig. 4.2). Specifically, with the

lack of CL (Δ*crd1*) or accumulation of MLCL (Δ*taz1*), the immunoblots against Tim23 and Tim17 (discussed below) showed a nearly complete absence of this complex. With the decreased association of Tim23 with the TIM23^{SORT*} assembly state, there was an increase in either the TIM23^{SORT} or the TIM23^{CORE} assembly states. Quantification of these analyses showed that for Tim23, Δ*taz1* mitochondria show a preferential accumulation of the TIM23^{SORT} assembly whereas in Δ*crd1* mitochondria there was an accumulation of the TIM23^{CORE} assembly state (Fig 4.3). These results suggest that the presence of MLCL (Δ*taz1*) or the lack of CL (Δ*crd1*) resulted in a destabilization of the TIM23^{SORT*} assembly state. When there was an accumulation of MLCL, the decrease in the TIM23^{SORT*} assembly state resulted in accumulation of the TIM23^{SORT} assembly state whereas when CL was completely absent the decrease in the TIM23^{SORT*} assembly state resulted in an increase in the TIM23^{CORE} assembly state. This suggests that in the absence of CL the TIM23^{SORT} complex is destabilized. 2D BN/SDS-PAGE analysis confirmed these patterns and additionally resolved that there was an increase in the amount of free Tim23 in Δ*crd1* mitochondria (Fig 4.4).

As mentioned above, there was a dramatic decrease in the association of Tim17 with the TIM23^{SORT*} assembly state in mitochondria from Δ*taz1* and Δ*crd1* yeast strains, as visualized by BN-PAGE (Fig 4.2, Fig 4.3). 2D BN/SDS-PAGE analysis confirmed these patterns and additionally resolved that there was an increase in the amount of free Tim17 protein in both Δ*taz1* and Δ*crd1* mitochondria (Fig 4.4). This again suggests that CL destabilizes the TIM23^{SORT} complex.

To further understand TIM23 assembly state dynamics, we looked at the distribution of Tim21 in the TIM23^{SORT} assembly states. Analysis with αTim21 confirmed

the above results and shows that in Δ*taz1* and Δ*crd1* mitochondria the TIM23^{SORT*} assembly state is destabilized (Fig 4.2). Quantification of the proportion of Tim21 in each assembly state showed there was a decrease of 30% and 35% in the proportion of Tim21 in the TIM23^{SORT*} assembly state for Δ*taz1* and Δ*crd1* mitochondria, respectively (Fig 4.3). These results suggest, that again, in the presence of MLCL or the complete absence of CL there was a destabilization of the higher molecular weight TIM23^{SORT*} assembly state. 2D BN/SDS-PAGE with αTim21 confirms the destabilization of the TIM23^{SORT*} assembly state (Fig 4.4).

Finally, we addressed the assembly status of the Tim50 receptor in the different CL backgrounds using the above approaches. Previous research has suggested that Tim50 resides in the TIM23^{SORT*} assembly state (37), however we find Tim50 is present in the lower molecular weight TIM23^{SORT} assembly state (Fig 4.2) and that there is little effect seen with Tim50 association with the TIM23^{SORT} assembly state.

These results suggest there is a CL dependent stabilization of the TIM23^{SORT} complex. Our results also show that TIM23 subunits can dynamically assemble into different complexes in a CL-dependent fashion. Moreover, this work suggests that TIM23^{CORE} may act as a 'repository' to store channel-forming subunits of the protein import complex when CL concentrations are limiting, or in CL-poor microenvironments of the IM. We propose that because the TIM23^{SORT} complex specifically mediates the integration of proteins into the IM, its reliance on CL for assembly may ensure that nascent membrane proteins are integrated into CL-rich regions of the IM to promote integration, topogenesis, or both. These results are supported by previous experimental

evidence that CL-rich liposomes are necessary for reconstitution of actively integrating TIM23^{SORT} (229).

4.4.2 Uncoupling of the Mitochondrial IM Contributes to the TIM23^{SORT} Complex Destabilization

Previous research has indicated that Δ*taz1* and Δ*crd1* yeast mitochondria show decreased membrane potential (20,242). CL is necessary for proper function of the electron transport chain (ETC) and is necessary for the stabilization of respiratory supercomplexes, both of which support the generation of the IM membrane potential (20,25,240,243,244). To assess the degree to which destabilization of the TIM23^{SORT}* complex in the Δ*crd1* and Δ*taz1* strains results specifically from membrane depolarization, WT mitochondria were titrated with CCCP prior to native gel analysis. CCCP is an uncoupler that dissipates the membrane potential and effectively disconnects the ETC from the formation of ATP.

1D BN-PAGE (Fig 4.5) and 2D BN/SDS-PAGE (Fig 4.6) were used to resolve the CCCP-dependent differences in the distribution of the TIM23 assembly states. With decreased IM potential, Tim23 showed only a slight increase in free protein (Fig 4.6), with no discernible changes in its assembly state distributions. Tim21 also showed no changes in the distributions of the TIM23 assemblies (Fig 4.5, Fig 4.6). Tim17, however, did show an increase in the TIM23^{CORE} assembly state with increasing concentrations of CCCP (Fig 4.5, Fig 4.6). Finally, there was not a change in Tim50 association with the TIM23^{SORT} assembly state but there was an increase in free Tim50 protein (Fig 4.5, Fig 4.6).

Together, these results show that although a reduced membrane potential may destabilize the interaction of Tim17 with the TIM23^{SORT} complex and increase the amount of unassembled Tim50, it did not recapitulate the same trends seen above with Δ*taz1* and Δ*crd1* mitochondria. We conclude that although the reduced membrane potential in Δ*taz1* and Δ*crd1* could be contributing to the altered proportions of the TIM23 complex it is not responsible for the overall destabilization of the TIM23^{SORT} complex that is seen in Δ*taz1* and Δ*crd1* mitochondria. Hence, properly remodeled CL species are required for TIM23^{SORT} assembly in a manner that goes beyond the contribution of CL to the energized status of the IM.

4.4.3 CL Supplementation Shifts the Distribution of the TIM23 Assembly States

To understand if supplemental CL could reverse the destabilizing effects seen on the TIM23 assembly states when CL is absent, we added CL-containing NDs and added them to Δ*crd1* mitochondria, and analyzed the samples by BN-PAGE (Fig 4.7), and performed western blot with αTim23. Preliminary data suggests a concentration dependent shift of the TIM23 assembly states when CL containing NDs were added (Fig 4.7). Specifically, when CL was added at a stoichiometric ratio of 10:1 relative to endogenous CL levels found in WT mitochondria (Fig 4.7, 10-Fold) there was a decrease in the amount of the TIM23^{CORE} assembly state and a slight increase in the TIM23^{SORT*} assembly state, that was not seen with untreated Δ*crd1* mitochondria. At twenty times the amount of CL added, there was a defined shift away from Tim23 association with the TIM23^{CORE} assembly state and an increase of Tim23 association with both of the TIM23^{SORT} assembly states (Fig 4.7, 20-Fold). Finally, when fifty times

the amount of CL was added there was a dramatic effect, in that the distribution of Tim23 shifted almost entirely to the TIM23^{SORT}* assembly state (Fig 4.7, 50-Fold).

These preliminary results provide two major takeaways that must be studied further. First, the destabilizing effects seen in Δ*crd1* mitochondria seem to be reversible in that the addition of supplemental CL can shift the distribution of Tim23 away from the TIM23^{CORE} assembly state and towards the TIM23^{SORT}assembly states that mimics the distributions seen in WT mitochondria. Second, these results suggest that the CL that has been assembled into NDs is able to traverse the OM and integrate into the IM to stabilize these IM proteins. This comprehensive study of the assembly states of the TIM23 complex can now be used to understand the effects of supplemental CL, not only on the assembly states of TIM23 but to provide insights into the potential application of using NDs as a lipid delivery system, which will be discussed in more detail below.

4.5 Conclusions and Future Directions

The modular nature of the TIM23 complex allows multiple functional organizations of the complex and the unique ability to both integrate proteins in the IM and translocate proteins into the matrix. This modular organization is necessary for the function of the TIM23 complex but also results in the fluid association of subunits dependent on complex environment and functional state. Here, we have shown that a lack of CL or the accumulation of MLCL in the mitochondria results in a destabilization of the TIM23^{SORT} complex and when CL is absent there is an accumulation of the Tim23^{CORE} assembly state. Interestingly, the lack of any strong phenotype associated with the Δ*cld1* mutants shows that nascent CL supports TIM23^{SORT} assembly as well as fully-remodeled CL. We therefore conclude that tetra-acyl CL species, regardless of

their specific acyl chain composition, are key to the assembly of $TIM23^{SORT}$. We also show that these alterations in the assembly states of TIM23 are not a strict result of the decrease in membrane potential seen in Δ*taz1* and Δ*crd1* yeast strains.

We have also shown preliminary data suggesting that the destabilization of the $TIM23^{SORT}$ complex due to a lack of CL is reversible with the addition of CL containing NDs. Future research must be done to validate these results and to gain understanding in how CL is able to cross the OM and integrate in the mitochondrial IM and stabilize these assembly states. This optimized BN-PAGE assay will provide an effective readout for understanding the effects of supplemental CL on the assembly states of the TIM23 complex. While NDs have traditionally been used to study membrane proteins, we have shown here that NDs may also have potential as a lipid-delivery system. Previous research suggest CL containing NDs localize to mitochondria but this research failed to provide a clear mechanism for understanding how or in what capacity the CL may have been integrated into mitochondria. This work serves as a starting point for our development of a ND-based CL delivery system that can be specifically targeted to the mitochondria through the addition of an N-terminal presequence to MSP (Fig 4.8A, B). By experimenting with the sequence of the presequence (245) attached to MSP we would be able to optimize ND targeting to the mitochondria (Fig 4.8C, left). Additionally, we hypothesize that once the presequence has been engaged by the TOM and TIM23 (Fig 4.8, middle) complexes an unraveling of MPS from the ND would result in the targeted release of the ND lipids in the vicinity of the mitochondria (Fig 4.8, right).

Together these results have provided a comprehensive understanding into the effect of CL on the TIM23 assembly states and have preliminarily shown that this effect

is reversible. These results open exciting avenues of research and the development of using NDs as a mitochondria targeted lipid delivery system that could have profound effects on heritable diseases involving defects in lipid biosynthesis, trafficking, and remodeling such as Barth Syndrome.

4.6 Figures

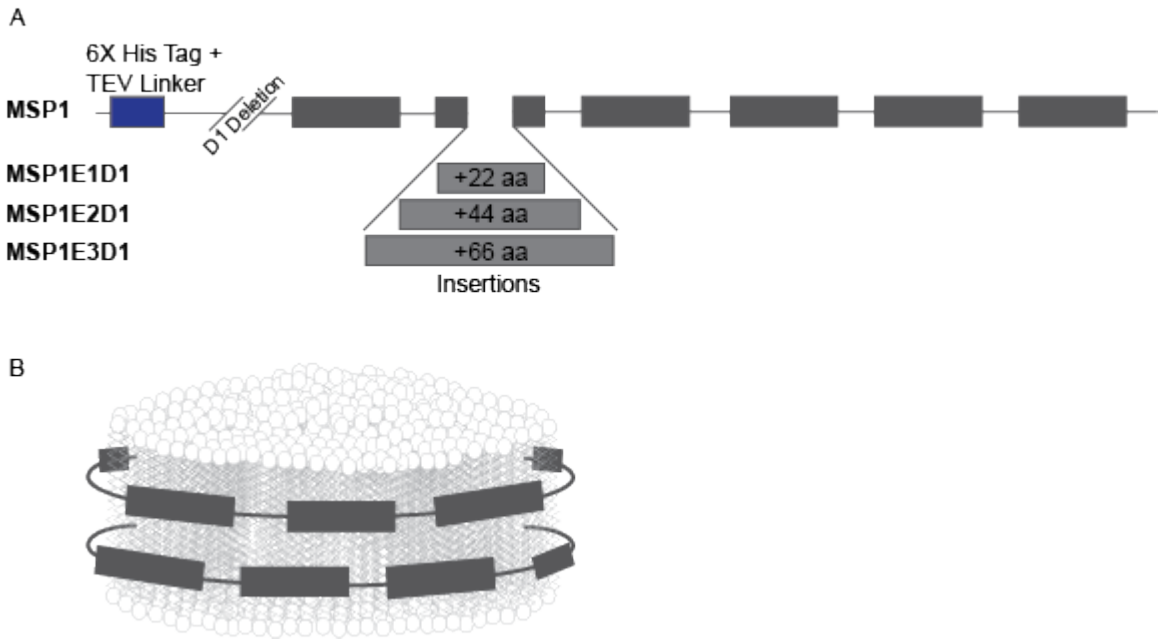


Figure 4.1. MSP is used for ND assembly. (A) Schematic of MSP. MSP1 is the simplest form of MSP with a hexa-histidine tag separated by a TEV cleavage site and linker. Multiple forms of MSP exist with a D1 deletion of 22 amino acids and an increasing number of 22-amino acid amphipathic helices denoted (MSP1E1D1-MSP1E3D1). (B) Two copies of MSP encircle the discoidal lipid bilayer to form a ND of desired lipid composition and diameter depending on the MSP form used.

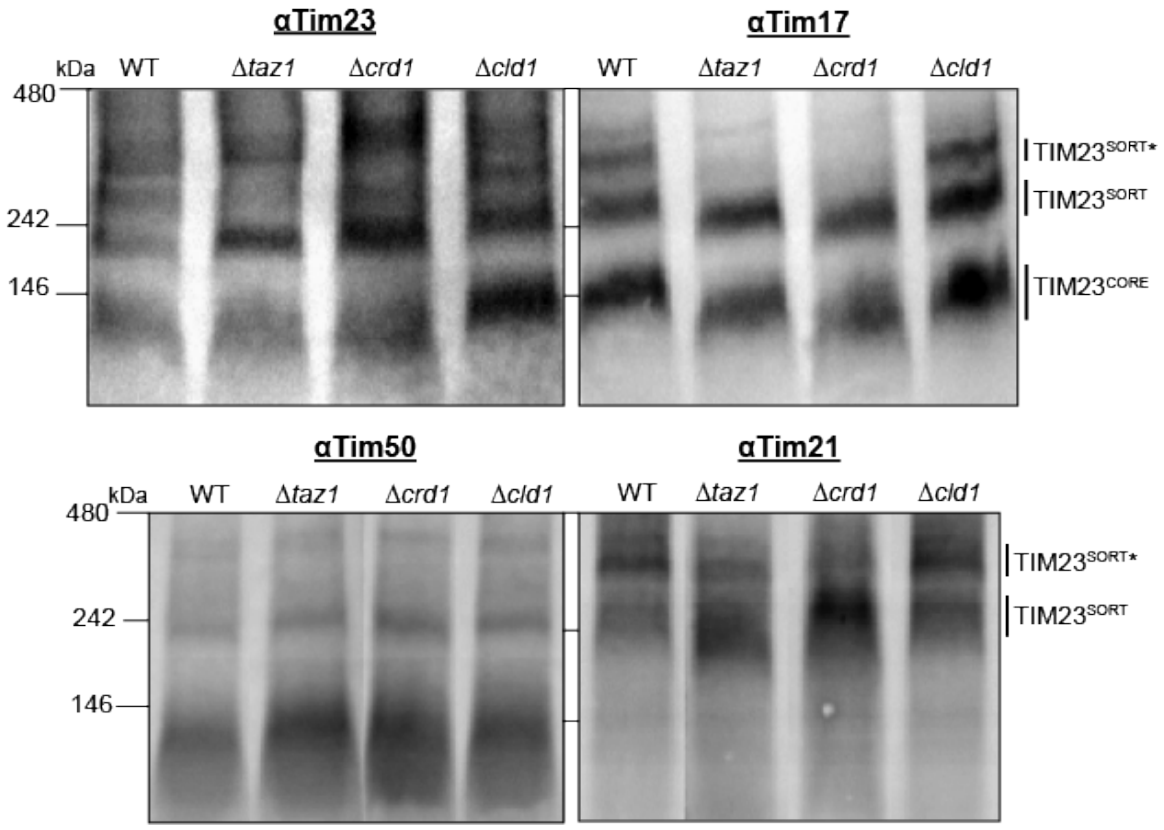


Figure 4.2. BN-PAGE analysis of the effects of CL on the assembly states of the TIM23 complex. Mitochondria from different CL backgrounds were solubilized with digitonin, separated by BN-PAGE, and analyzed by western blot analyses against Tim23, Tim17, Tim21, or Tim50 as indicated. The positions of the molecular weight markers are indicated at the left side of the figure. Bold lines at the right side of the figure indicate the identified TIM23 assembly states.

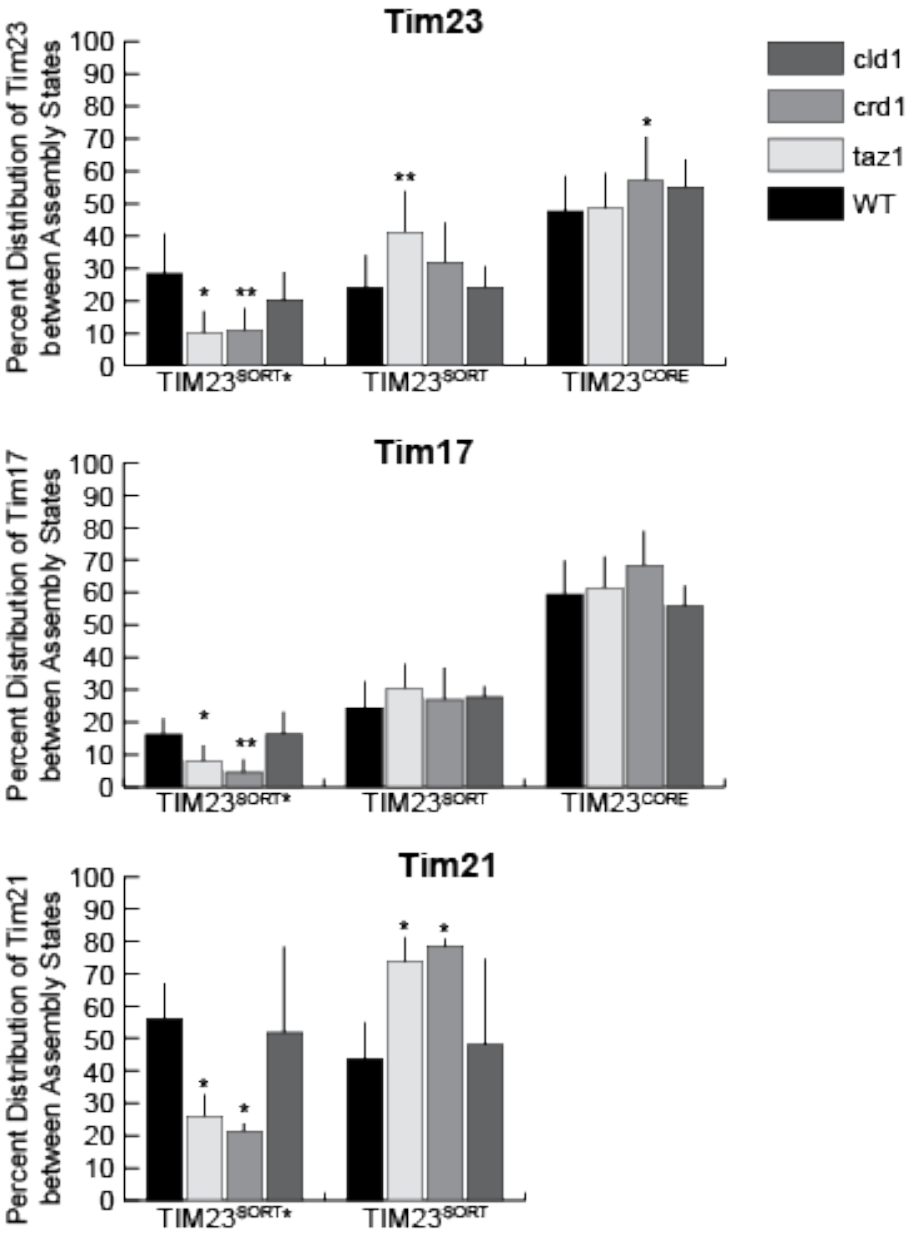


Figure 4.3. Quantification of the assembly states of the TIM23 complex. *Top*, Tim23. Middle, Tim17. Data are mean ± standard deviation of five independent experiments. Bottom, Tim21, data are mean ± standard deviation of three independent experiments. For each protein analyzed, statistical analyses were done for each assembly state and compared each strain to WT.

* Indicates P< 0.1. **Indicates P<0.01.

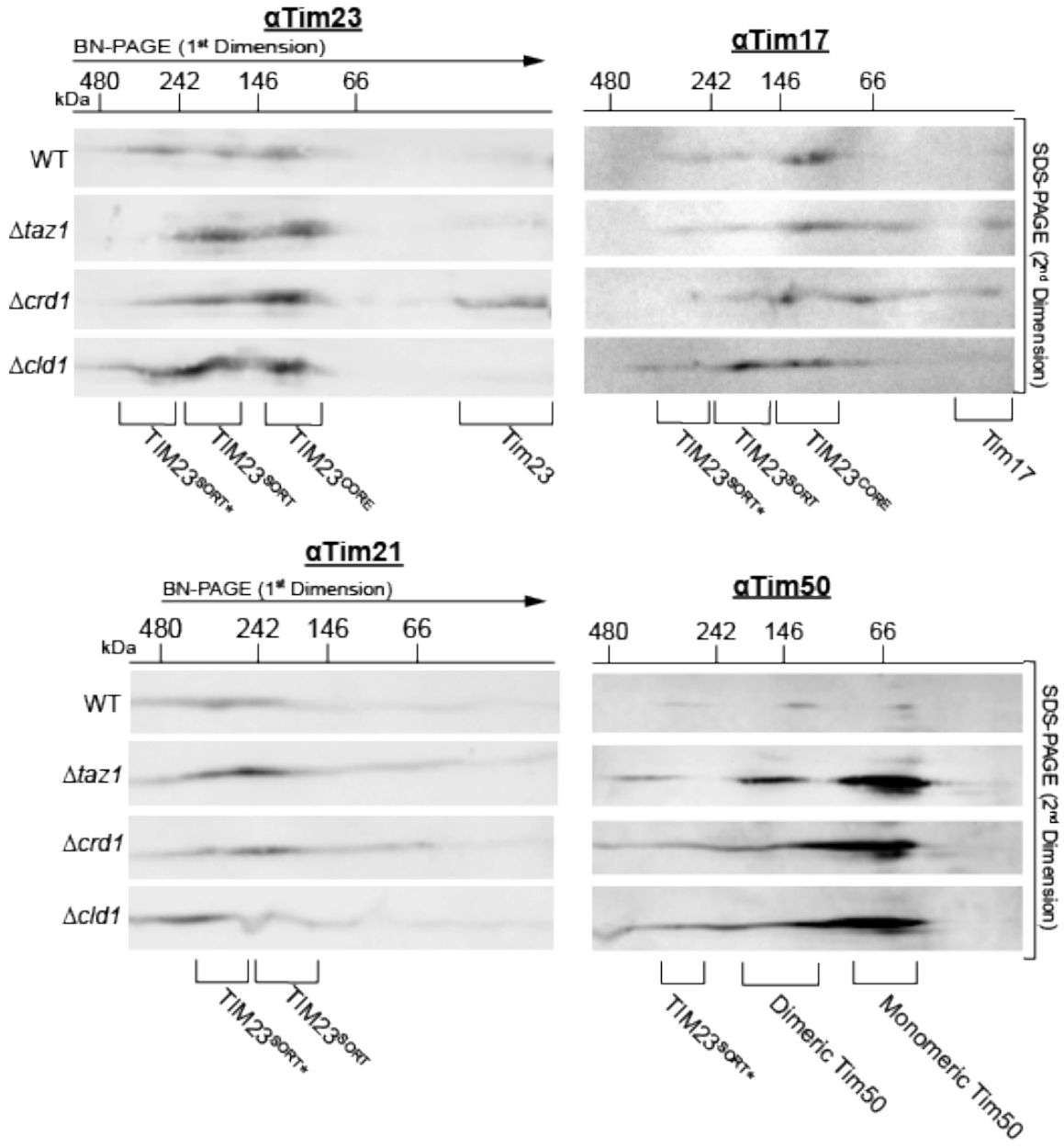


Figure 4.4. 2D BN/SDS-PAGE analysis of the effects of CL on the assembly states of the TIM23 complex. Mitochondria from different CL backgrounds (as indicated at the left side of the figure) were solubilized with digitonin and separated by BN-PAGE (1st dimension). Each sample lane was excised and additionally resolved by SDS-PAGE (2nd dimension) and analyzed by western blot analyses

against Tim23, Tim17, Tim21, or Tim50 as indicated. The positions of the molecular weight markers are indicated in kDa on the top of each set of images. Assembly states of the TIM23 complex are indicated below each set of images.

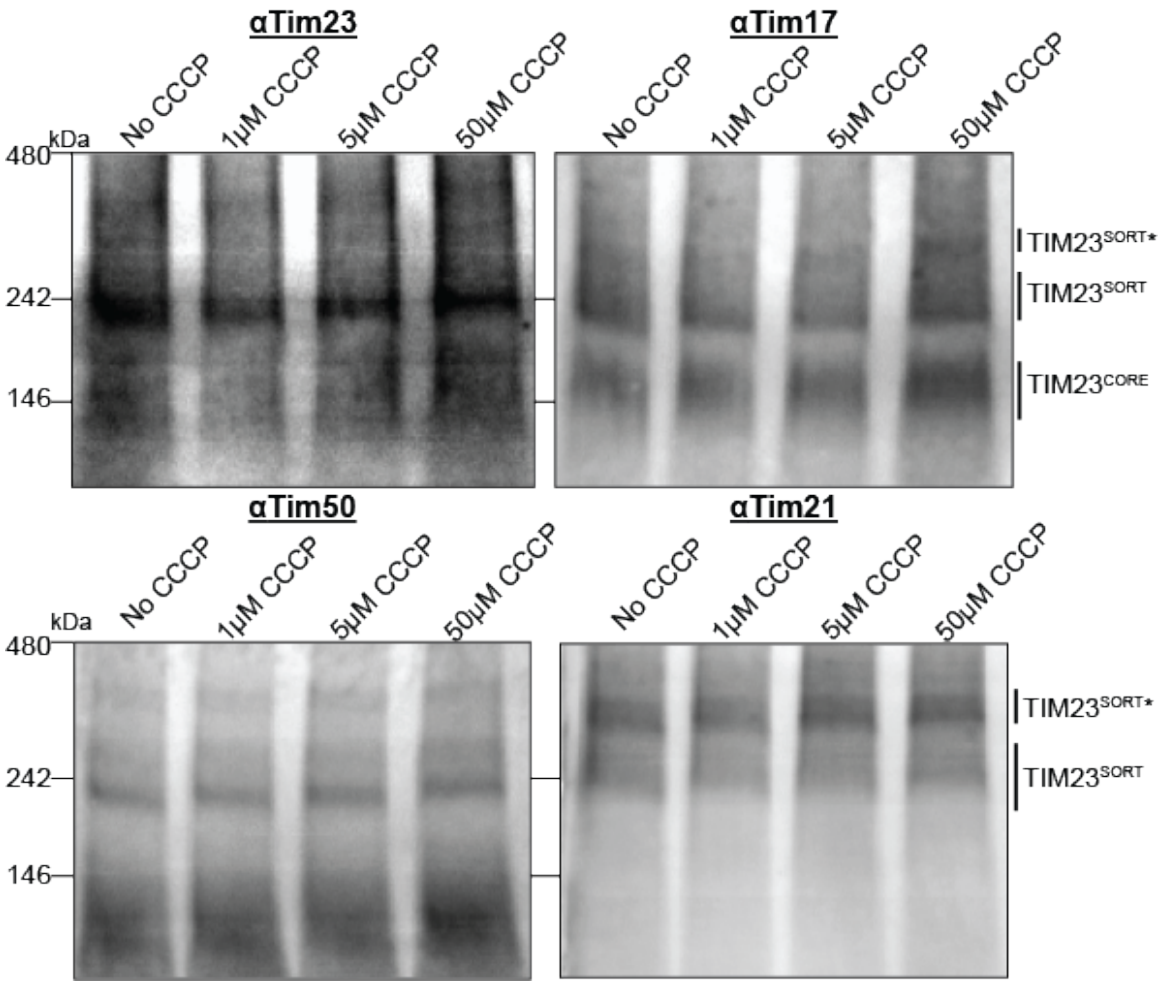


Figure 4.5. BN-PAGE analysis of the effects of dissipated membrane potential on the TIM23 assembly states. WT mitochondria were treated with varying concentrations of CCCP (as indicated at the top of each image), solubilized with digitonin, separated by BN-PAGE, and analyzed by western blot against either Tim23, Tim17 or Tim21, as indicated. The positions of the molecular weight markers are indicated at the left side of the figure. Bold lines at the right side of the figure indicate the identified TIM23 assembly states.

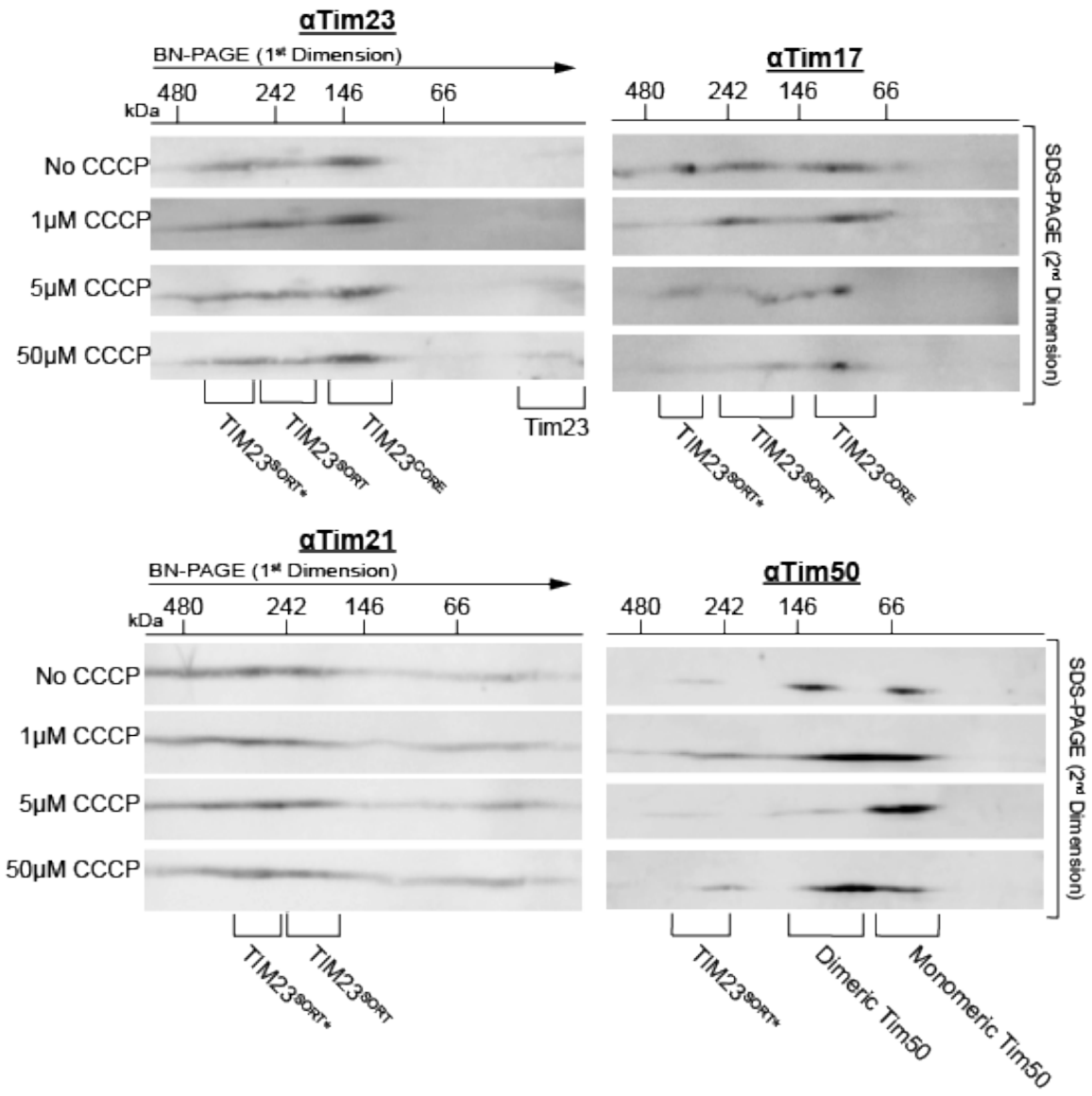


Figure 4.6. 2D BN/SDS-PAGE analysis of the effects of dissipated membrane potential on the TIM23 assembly states. WT mitochondria were treated with varying concentrations of CCCP (as indicated at the left side of the figure) and then solubilized with digitonin and separated by BN-PAGE (1st dimension). Each sample lane was excised and additionally resolved by SDS-PAGE (2nd dimension) and

analyzed by western blot analysis with the indicated antibody. The positions of the molecular weight markers are indicated in kDa on the top of each set of images. Assembly states of the TIM23 complex are indicated below each set of images.

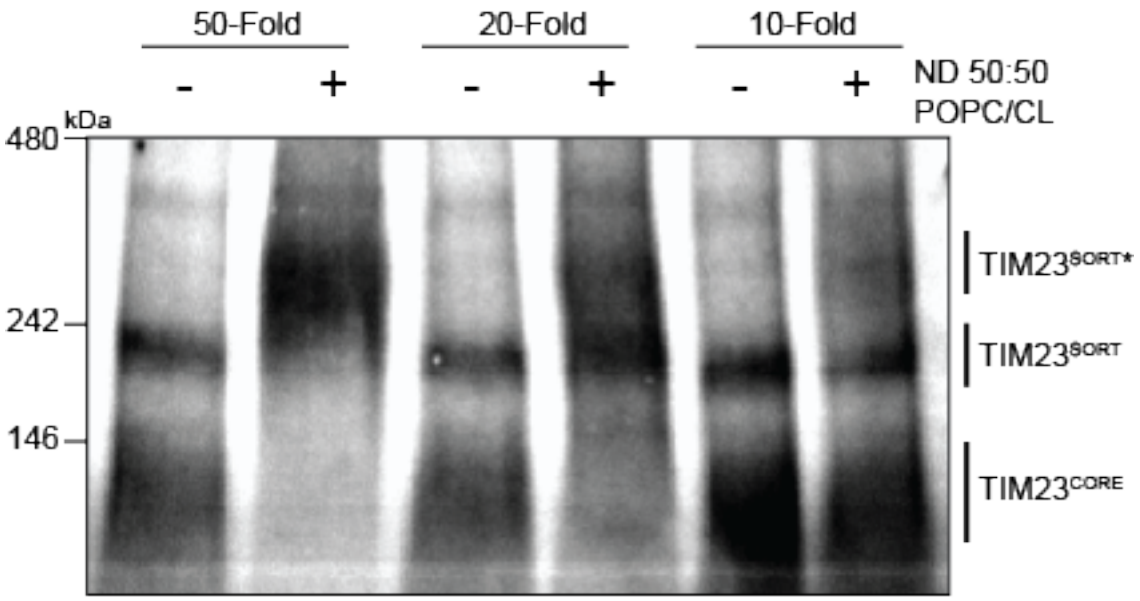


Figure 4.7. CL supplementation shifts the distribution of TIM23 assembly states. Δ*crd1* mitochondria were supplemented with varying concentrations of CL containing NDs. CL was introduced to isolated mitochondria at a ratio of 10, 20, or 50 times (10-Fold, 20-Fold, 50-Fold) the amount of CL found in WT IM membranes. Molecular weight markers are denoted on the left side of the image and identified TIM23 assembly states are denoted on the right.

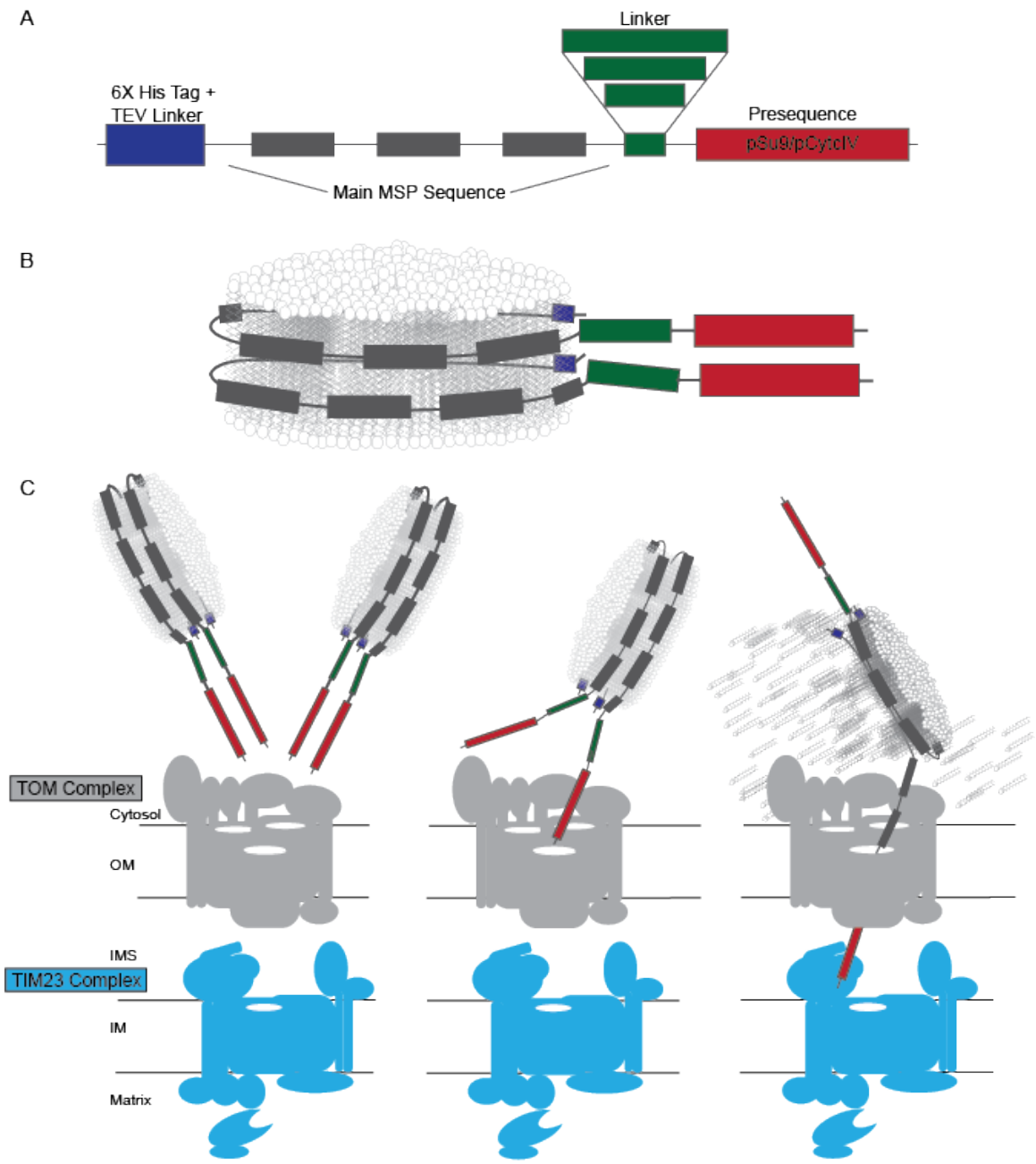


Figure 4.8. NDs as a mitochondria targeted lipid delivery system. (A) Schematic of an MSP design to integrate a mitochondrial presequence. Optimization of the linker region and presequence will allow for efficient targeting of MSP to the

mitochondria. Possible presequences include the presequence of subunit nine of the F_o-ATPase (pSu9), targeted to the mitochondrial matrix, or the presequence of subunit four of cytochrome c oxidase (pCytcIV), targeted to the IM. (B) Schematic of an assembled ND with MSP including a mitochondrial presequence. (C) *Left*, NDs formed using MSP with an integrated presequence will be targeted to the mitochondria, specifically through the TIM23 pathway. *Middle*, upon interaction with the TOM and TIM23 complexes import of MSP will result in, *right*, the unraveling of MSP and release of ND lipids.

Chapter 5

Effect of *N*-acetyl-*p*-benzoquinone imine (NAPQI) on Respiratory Complex II

Derived from:

Lee KK. Imaizumi N., **Chamberland SR**., Alder NN., Boelsterli UA. (2015) Targeting Mitochondria with Methylene Blue Protects Mice Against Acetaminophen-Induced Liver Injury *Hepatology* **61**, 326-336.

***Collaborative work**

5.1 Abstract

Acetaminophen (APAP) overdose is the most frequent cause of drug-induced liver injury in the Western world. Previous studies with mouse models have revealed that impairment of mitochondrial respiration is an early event in the pathogenesis, but the exact mechanisms have remained unclear, and therapeutic approaches to specifically target mitochondria have been insufficiently explored. Here, we found that the reactive oxidative metabolite of APAP, *N*-acetyl-*p*-benzoquinone imine (NAPQI), caused the selective inhibition of mitochondrial complex II activity by >90% in both yeast-derived complexes reconstituted into nanoscale model membranes and in mouse hepatic mitochondria, as well as the decrease of succinate-driven ATP biosynthesis rates. Based on these findings we hypothesized that methylene blue (MB), a mitochondria-permeant compound that can act as an alternative electron carrier, protects against APAP-induced hepatocyte injury. We found that MB ($\leq$ 3 mM) readily accepted electrons from NAPQI-altered, succinate-energized complex II and transferred them to cytochrome *c*, restoring ATP biosynthesis rates. In conclusion, MB can effectively protect mice against APAP-induced liver injury by bypassing the NAPQI-altered mitochondrial complex II, thus alleviating the cellular energy crisis. Because MB is a clinically used drug, its potential application after APAP overdose in patients should be further explored.

5.2 Introduction

5.2.1 The OXPHOS System

The majority of ATP synthesis occurs at the mitochondrial IM through oxidative phosphorylation (OXPHOS). The OXPHOS system is composed of proteins of the

electron transport chain (ETC), specifically, respiratory complexes I-IV along with the F_1F_o-ATP synthase, which is also known as complex V. Upon passage of electrons through complexes I-IV, energy is generated in the form of a proton gradient across the IM (Fig 5.1). The energy of this proton gradient is subsequently used to drive the conversion of ADP to ATP by the F_1F_o-ATP synthase. Complex I, known as NADH: ubiquinone oxidoreductase, is only found in higher eukaryotes. In yeast, complex I is replaced by multiple NADH dehydrogenases. In both systems, these proteins transport protons across the membrane, aiding in the generation of the membrane potential. Succinate: ubiquinone oxidoreductase (complex II) transports electrons from succinate to ubiquinone, and will be discussed in more detail below. Cytochrome bc_1 (complex III) oxidizes ubiquinol, resulting in the transfer of electrons to cytochrome *c* and protons being pumped to the IMS. Cytochrome *c* oxidase (complex IV), then utilizes electrons donated from cytochrome *c* to reduce O_2 into two molecules of H_2O that again results in protons being pumped into the IMS. Finally, F_1F_o-ATP synthase (Complex V) produces ATP by utilizing the energy stored in the proton gradient created by complexes I-IV (24,246).

5.2.2 Complex II Structure and Function

Respiratory complex II is heterotetrameric in nature (Fig 5.2) (247-249). The soluble catalytic dimer is located in the mitochondrial matrix and is anchored to the IM by an integral membrane dimer. The catalytic dimer oxidizes succinate to fumarate as part of the tricarboxylic acid (TCA) cycle, referred to as the succinate dehydrogenase (SDH) activity. The catalytic dimer consists of subunits, Sdh1p and Sdh2p in yeast, and SdhA and SdhB in mammals. Sdh1p houses a covalently bound FAD cofactor, which

allows the oxidation of succinate. Sdh2p contains three disparate iron-sulfur clusters necessary for electron transport from FAD to ubiquinone. The integral IM dimer consists of subunits Sdh3p and Sdh4p in yeast, and SdhC and SdhD in mammals,that coordinate a *b*-type heme and contain the quinone-binding site (250,251). The transfer of electrons from succinate to ubiquinone is referred to as the succinate: ubiquinone oxidoreductase (SQR) activity of complex II.

5.2.3 Acetaminophen Induced Liver Injury

Acetaminophen (APAP) is the most widely used painkiller in the United States. However, when taken in high doses, either intentionally or unintentionally, APAP can lead to liver injury and eventually acute liver failure (252). While this hepatotoxicity is common, the signaling pathway is not entirely understood. What is known has mostly been investigated in mouse models (253-255) and shows mitochondria are critically important in this process. The oxidative metabolite of APAP, *N*-acetyl-*p*-benzoquinone imine (NAPQI), has been shown to interact with nucleophilic targets resulting in oxidative stress and then, in subsequent phases, cell death (256-258). Once NAPQI has been formed in the cytosol or in the mitochondria, it avidly reacts with multiple targets, including cysteine residues on mitochondrial proteins (259-261). NAPQI can also interfere, by poorly understood mechanisms, with the ETC. Previous studies have shown that APAP, and as a result NAPQI, inhibits both complex I and II without altering complex III and IV activity (262-265).

5.3 Materials and Methods

5.3.1 Preparation of Yeast Complex II Enzymes Reconstituted into Nanoscale Lipid Bilayers (Nanodiscs) and Measurement of Succinate:Ubiquone Oxidoreductase Activity Toxicity

Complex II derived from *S. cerevisiae* (strain D273-10B) was reconstituted into nanodiscs (Section 4.2.2.) as described (235). Briefly, isolated mitochondria were solubilized in buffer containing 2% (w/v) dodecyl-β-D-maltoside, and clarified supernatants containing complex II were used for nanodisc assembly reactions with MSP1E3D1 and phospholipids POPC, POPE and tetralinoleoyl CL in molar ratios of 2:2:1. Samples were subjected to five-minute preincubation steps, first with a range of NAPQI concentrations (up to 100 μM), and then with 10 μM decylubiquinone, before addition of assay buffer (50 mM potassium phosphate, pH 7.5, 1 mM EDTA, 62 μM DCPIP). SQR activity was initiated by the addition of 20 mM succinate and monitored over two minutes by measuring the extinction of DCPIP absorbance at ten second intervals in an Amersham Biosciences Ultraspec 2100 pro UV/Vis spectrophotometer (Fig 5.3). Measurements over the linear response range were used to calculate relative SQR activity.

5.4 Results and Discussion

5.4.1 MSP Purification

MSP was purified as described (266). BL21 *Escherichia coli* cells were transformed with the pET28a-MSP1E3D1 vector, grown to an OD_{600} of approximately 0.8, induced with isopropyl β-D-1-triogalactopyranoside, lysed, and collected by centrifugation. A sample of supernatant and pellet was collected. Purification of MSP

occurred through affinity chromatography with Ni-NTA agarose and increasing concentrations of imidazole. The quality of the purification was resolved by SDS-PAGE (Fig 5.4) and informed the elution fractions taken for dialysis and quantification. After dialysis, necessary to remove residual imidazole, spectroscopic quantification of MSP was done using A_{280} values.

5.4.2 NAPQI Inhibits Yeast Complex II Reconstituted into Nanodiscs

To confirm the inhibitory effects of NAPQI on complex II enzymatic activity, yeast mitochondria were reconstituted into stable nanodiscs of defined lipid composition (235). Following exposure to increasing concentrations of NAPQI, reconstituted complex II showed a marked reduction in SQR activity, with >90% inhibition with 20μM NAPQI (Fig 5.5). This reductionist nanodisc system was necessary in confirming that complex II is specifically inhibited by NAPQI. The use of the complex II activity assay allowed a clear understanding of the effect of NAPQI on SQR activity.

5.5 Conclusions

The inhibition of complex II activity by NAPQI was confirmed by our collaborators (Lee KK., Imaizumi N., Boelsterli UA.) in mice mitochondria. With this inhibition of complex II determined, our collaborators continued this study to further test if MB could bypass the NAPQI inhibition of complex II. MB is an alternative electron carrier that is currently used as a redox-active drug and importantly has been shown to bypass compromised complex I (267). Previous research has shown that MB can bypass complex I by accepting electrons from NADH at the NADH-binding site on complex I and transferring them to more distal sites of the ETC. This study expanded this research

by identifying that MB can circumvent complex II inhibition by NAPQI by transferring electrons from succinate to more distal sites of the ETC. Our collaborators showed that there were decreased ATP biosynthesis rates of mouse mitochondria with treatment with NAPQI. However, when mitochondria were treated with both NAPQI and MB there was a rescue of ATP biosynthesis rates. Our collaborators also looked at the effect of NAPQI and MB on liver damage in mice. Monitoring alanine aminotransferase (ALT) in blood is a read out for liver damage. When mice were treated with APAP there was a significant increase in ALT levels indicating liver damage, but when mice were treated with both APAP and MB, there was a return to near normal levels of ALT, meaning the extensive liver damage caused by APAP alone was attenuated by MB. These results allowed for a mechanism of action for NAPQI inhibition of complex II and attenuation by MB (Fig 5.6).

Currently, N-acetylcysteine is used to prevent the progression of acute liver failure caused by APAP overdoses. However, this treatment must be given within a relatively short time window (268). Here, we identified a possible alternative treatment for APAP induced acute liver failure through the use of MB, an already clinically used drug.

This study focused on the effects of NAPQI and MB on complex II activity within the larger process of APAP induced liver injury. The dysfunction of complex II in the presence of NAPQI results in the interruption of the inter-protein communication of the ETC. The reconstitution of complex II into nanodiscs in this study allowed for the direct investigation of the effect of NAPQI on complex II using a complex-specific SQR activity

assay that was a necessary basis for understanding the mechanism of APAP induced acute liver failure.

5.6 Figures

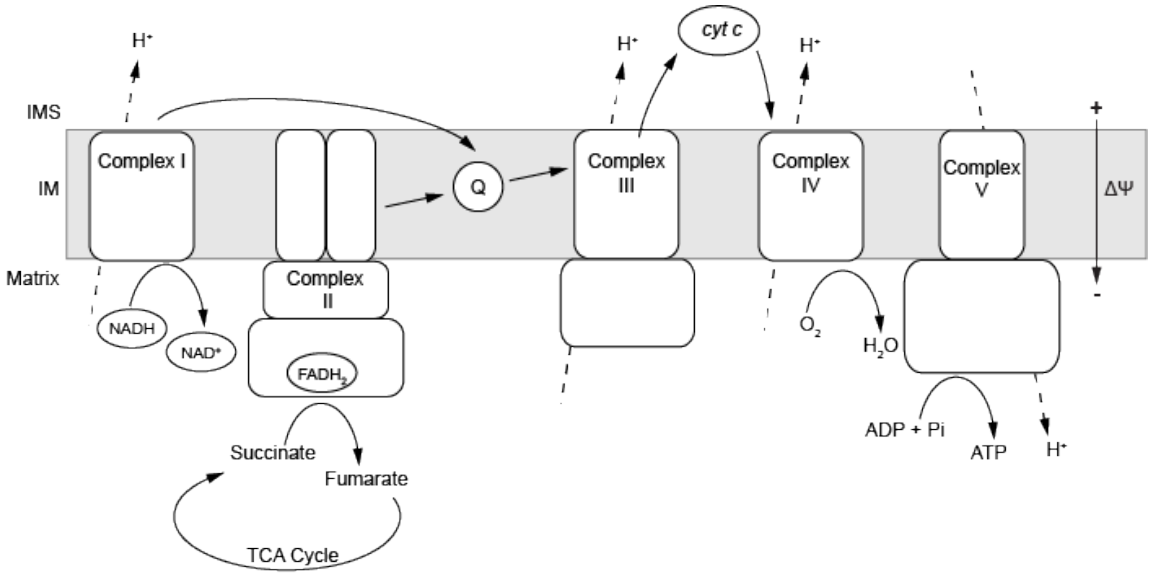


Figure 5.1. Schematic of the ETC of the mitochondrial IM. The ETC is located in the mitochondrial IM. NADH is utilized by complex I resulting in electrons being passed to ubiquinone (Q) and protons being transferred across the IM into the IMS. Complex II passes electrons from succinate to ubiquinone. Complex III transfers electrons from ubiquinone to cytochrome *c* (*cyt c*) in the IMS while contributing to the proton gradient. Electrons continue to complex IV resulting in the reduction of O_2 to H_2O and the pumping of protons into the IMS. Together this results in the formation of a proton gradient across the IM. Complex V utilizes the energy of this proton gradient to convert ADP and P_i into ATP in the mitochondrial matrix.

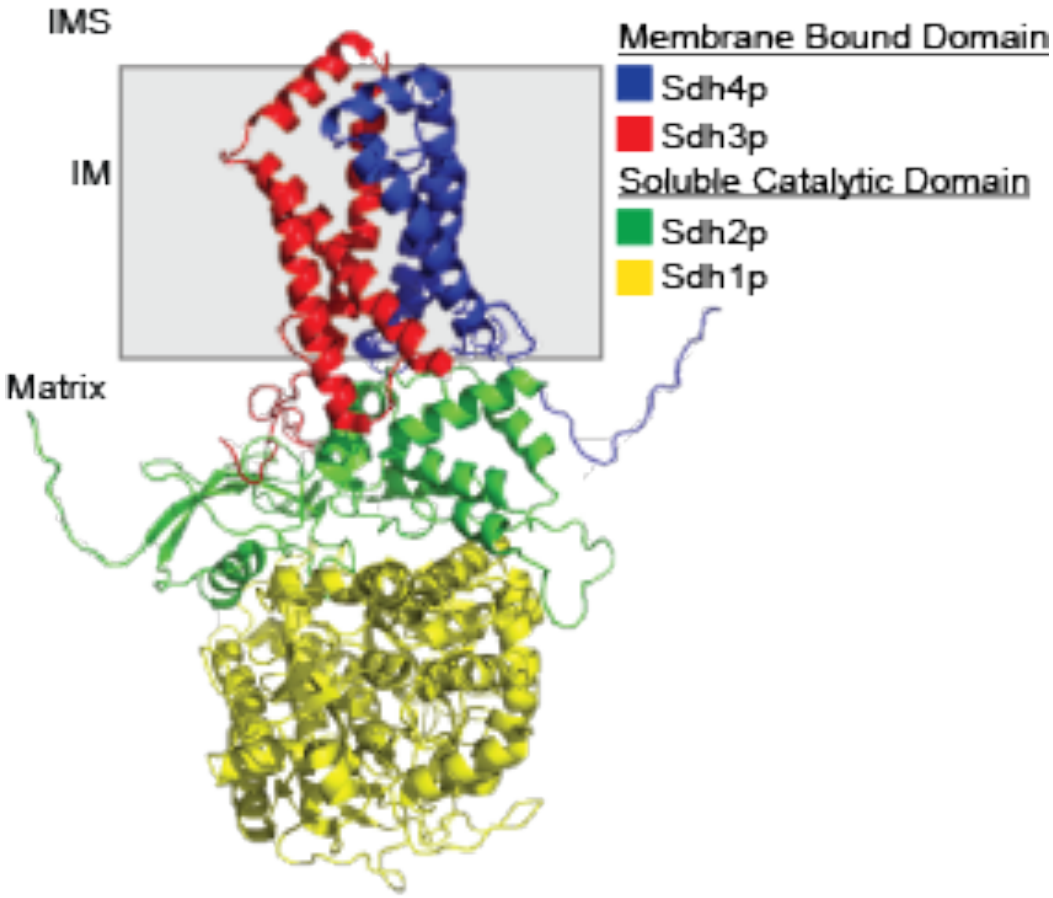


Figure 5.2. Homology model of yeast respiratory complex II. Figure is derived from the PDB structure 1PB4 (249). Complex II consists of a soluble catalytic dimer and membrane bound dimer. The catalytic dimer is composed of Sdh1p (yellow) and Sdh2p (green) located in the mitochondrial matrix. The membrane bound heterodimer is situated in the IM and composed of Sdh3p (red) and Sdh4p (blue).

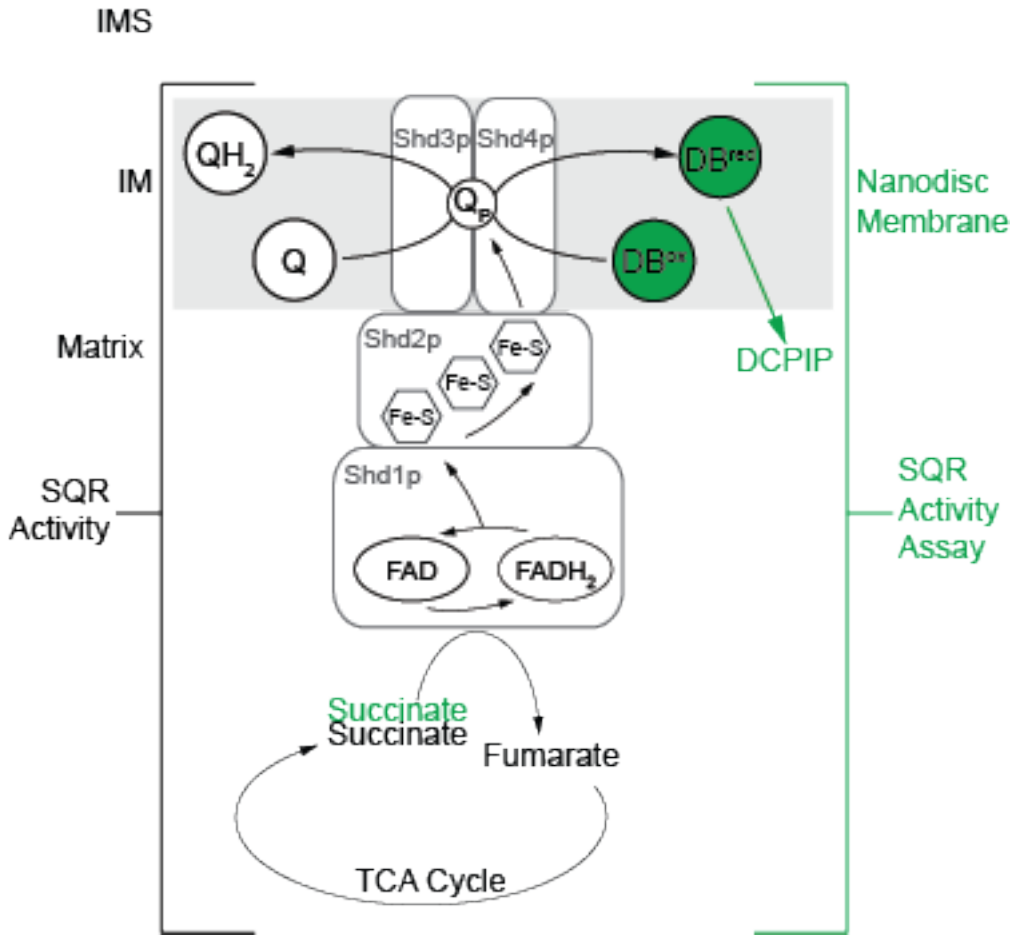


Figure 5.3. Complex II activity and the complex II SQR activity assay.

Complex II is part of both the ETC and TCA cycle. As part of the ETC, complex II transfers electrons from succinate to ubiquinone through its SQR activity. Electrons are transferred through FAD, the three iron-sulfur clusters (Fe-S), and then to ubiquinone (Q) (left side). The SQR activity assay utilizes this pathway with the addition of succinate and the ubiquinone analog, decylubiquinone (DB). Upon reduction of DB, electrons can then be passed to the artificial electron acceptor dichlorophenolindophenol (DCPIP), the reduction of which acts as a readout for SQR activity, which can be spectroscopically monitored at 600nm.

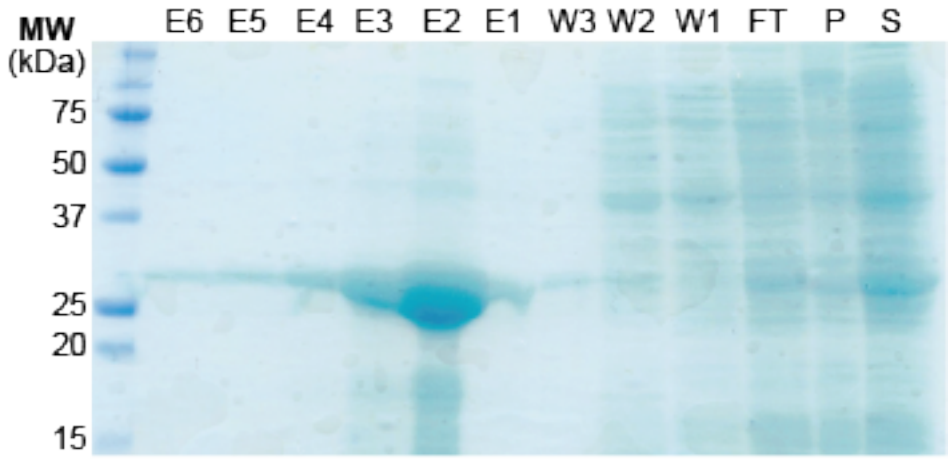


Figure 5.4. Purification of MSP. The MSP purification fractions were resolved by SDS-PAGE and coomassie staining. S, supernatant from *E. coli* lysate. P, pellet from *E. coli* lysate. An Ni-NTA column was used for purification. FT, column flow through. W1-W3, column wash steps. E1-E6, column elution steps. Imidazole concentrations were as follows, W1: 0mM, W2: 20mM, W3: 50mM E1-E6: 400mM.

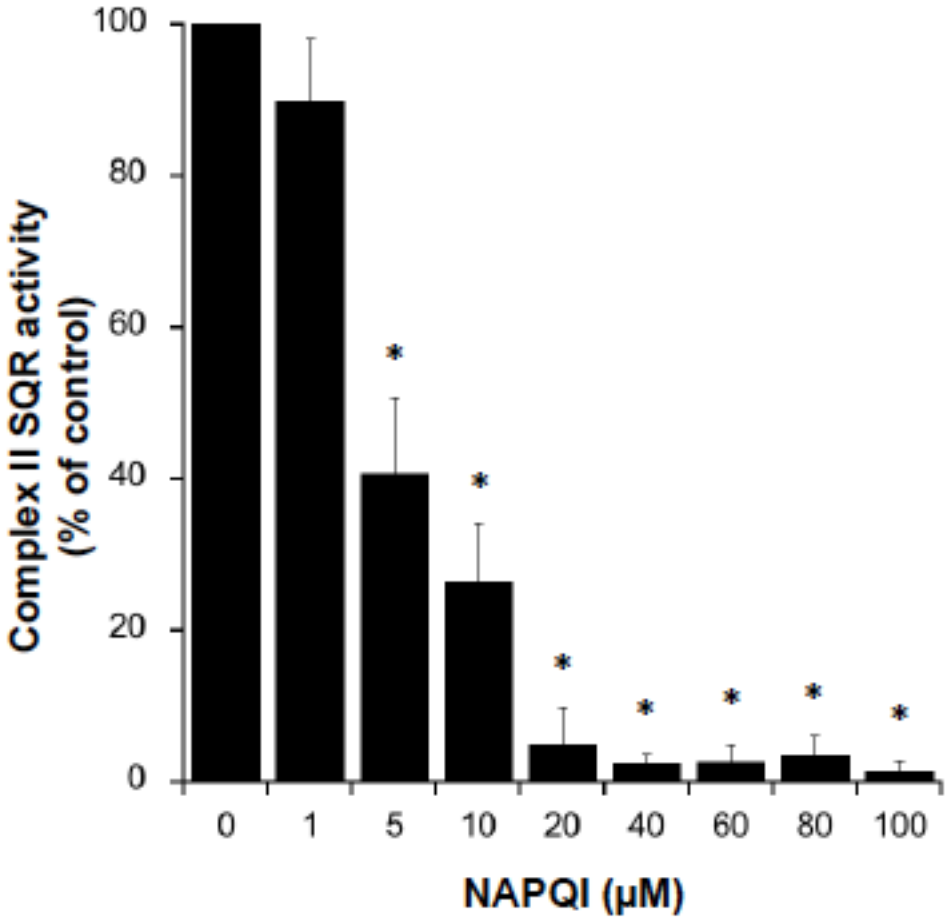


Figure 5.5. Effect of NAPQI on complex II SQR activity. SQR activity of yeast complex II reconstituted into nanodiscs following preincubation with various concentrations of NAPQI. Data are mean ± standard deviation of four independent nanodisc preparations. * indicates $P < 0.05$. 20μM NAPQI resulted in >90% inhibition of SQR activity.

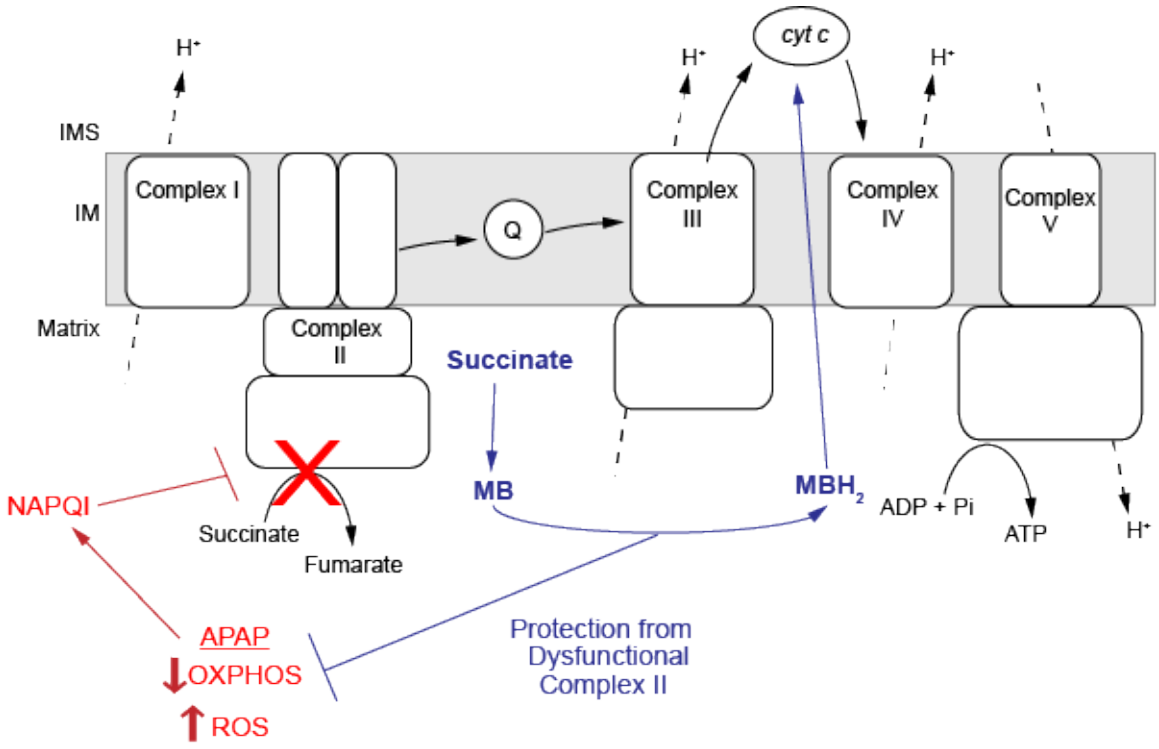


Figure 5.6. Schematic representation of the mechanism of MB bypassing compromised complex II. Upon APAP induced injury there is production of the oxidative metabolite NAPQI, a decrease in OXPHOS, and an increase in ROS production. MB can accept electrons from succinate when complex II is inhibited by NAPQI and unable to accept these electrons. MB then transfers electrons to to more distal sites of the ETC therefore restoring ETC function and protecting from dysfunction complex II and fighting against APAP induced injury.

Appendices

Appendix A – Coevolution Analysis of the Mitochondrial Carrier Family Provides Insights into Conformational Dynamics and Transport Mechanisms – Supplemental Figures

Appendix B – The TIM23 and TOM Complex Subunit Sequence Database Summary

Appendix A – Coevolution Analysis of the Mitochondrial Carrier Family Provides Insights into Conformational Dynamics and Transport Mechanisms – Supplemental Figures

	H1	m1	h12	m2	H2	C23	H3	m3	h34	m4	H4	C45	H5	m5	h56	m6	H6
H1	23																
m1	6	1															
h12	8	8	15														
m2	4		7	5													
H2	12			1	21												
C23					3	4											
H3	4				18	4	16										
m3	5						1	2									
h34					1		6	10	14								
m4							1		8	3							
H4	1				6		18				12						
C45																	
H5	3				1		6	2			18		14				
m5	1												2	1			
h56											2		6	11	16		
m6	2												6		7	7	
H6	12		1		2	2					3		14			2	9

Table A.1. Number of coevolving pairs sorted by structural region. Table includes the number of coevolving residue pairs within and between each structural element of yeast AAC (4C9G). All coevolving residue pairs are included in this table. H – Transmembrane helix. h - matrix-helix. m – matrix exposed loops. C – cytosolic exposed loops.

A

1st Predicted Residue	2nd Predicted Residue	Distance Catagory		
		4C9G	Model	6GCI
Lys 48/40	Gln 52/44			
Glu 45/37	Ala 156/149			
Arg 152/145	Ala 156/149			
Arg 46/38	Ala 156/149			
Asp 149/142	Arg 252/245			
Asp 149/142	Asp 249/242			

B

1st Predicted Residue	2nd Predicted Residue	Distance Catagory		
		4C9G	Model	6GCI
Lys 108/ Arg 100	Asp 212/205			
Asp 109/101	Gln 309/Lys 302			
Lys 112/104	Asp 212/205			
Tyr 211/204	Asp 306/299			

Table A.2. Coevolving pairs of the Matrix and Cytosolic Gates. (A) Predicted coevolving pairs of the c-state salt bridge and glutamine brace residues. Colored circles match the color of the residues in Figure 2.11. (B) Predicted coevolving pairs of the m-state salt bridge and tyrosine brace residues. Colored circles match the color of the residues in Figure 2.12. Distance category circles – green, <5Å apart; yellow, 5-10Å apart; red, >10Å apart.

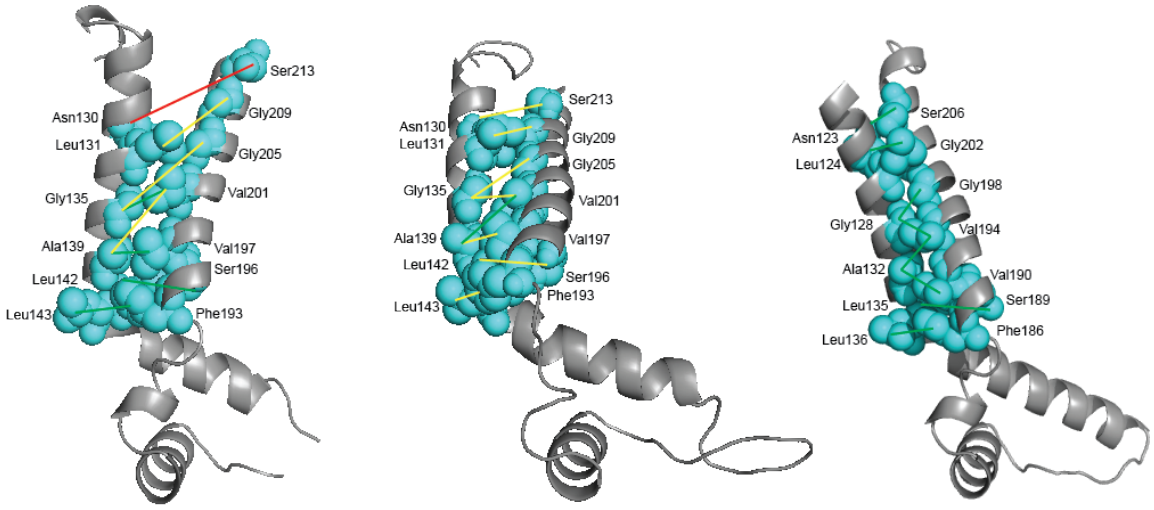


Figure A.1. Intramodule transmembrane helical packing of module two. The coevolution analysis identified residues in a classical i, i+4 pattern along one helical face, shown as colored spheres. These faces suggest the intramodule interaction of the transmembrane helices. *Left*, c-state structure. *Middle*, RaptorX Contact Model. *Right*, m-state structure. Distances between residues are depicted as color-coded lines (green, <5Å apart; yellow, 5-10Å apart; red, >10Å apart).

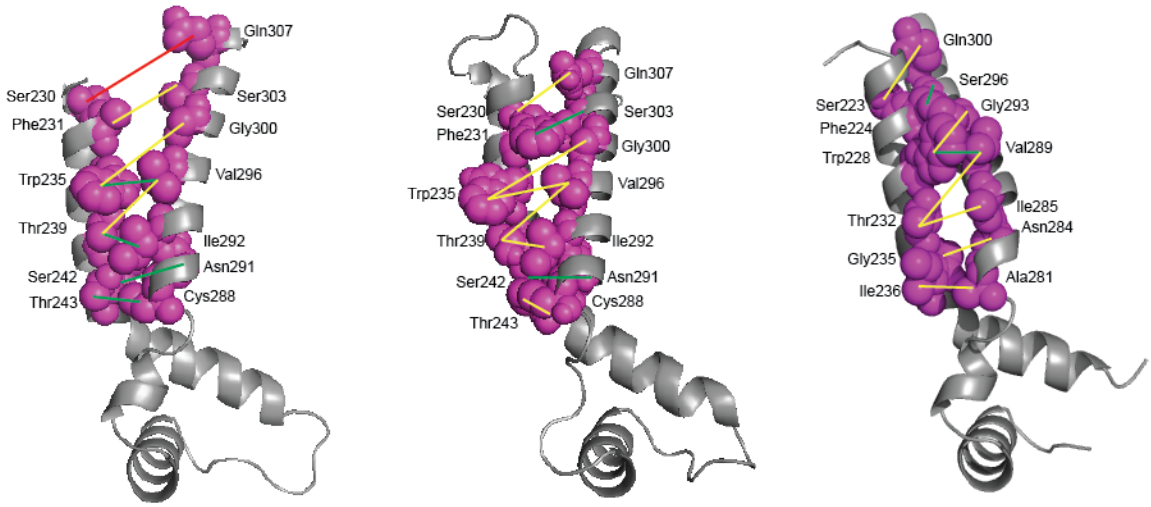


Figure A.2. Intramodule transmembrane helical packing of module three. The coevolution analysis identified residues in a classical i, i+4 pattern along one helical face, shown as colored spheres. These faces suggest the intramodule interaction of the transmembrane helices. *Left*, c-state structure. *Middle*, RaptorX Contact Model. *Right*, m-state structure. Distances between residues are depicted as color-coded lines (green, <5Å apart; yellow, 5-10Å apart; red, >10Å apart).

		Distance Catagory		
1st Predicted Residue	2nd Predicted Residue	4C9G	Model	6GCI
Asp 26/18	Asp 109/ Asp 101	🔴	🟢	🟢
Asn 130/123	Ser 213/206	🔴	🟡	🟢
Ser 230/223	Gln 307/300	🔴	🟡	🟡
Phe 27/19	Ala 106/98	🟡	🟡	🟢
Leu 131/124	Gly 209/202	🟡	🟡	🟢
Phe 231/224	Ser 303/296	🟡	🟢	🟢
Gly 31/23	Ala 102/94	🟡	🟡	🟢
Gly 135/128	Gly 205/198	🟡	🟡	🟢
Trp 235/228	Gly 300/293	🟡	🟡	🟡
Gly 31/23	Phe 98/90	🟢	🟢	🟢
Gly 135/128	Val 201/194	🟢	🟡	🟢
Trp 235/228	Val 296/289	🟢	🟡	🟢
Ala 35/27	Phe 98/90	🟡	🟡	🟢
Ala 139/132	Val 201/194	🟡	🟢	🟢
Thr 239/232	Val 296/289	🟡	🟡	🟡
Ala 35/27	Val 94/86	🟢	🟡	🟢
Ala 139/132	Val 197/190	🟢	🟡	🟢
Thr 239/232	Ile 292/285	🟢	🟡	🟡
Lys 38/30	Asn 93/85	🟢	🟢	🟢
Leu 142/135	Ser 196/189	🟢	🟡	🟢
Ser 242/ Gly 235	Asn 291/284	🟢	🟢	🟡
Thr 39/31	Asn 90/82	🟢	🟡	🟢
Leu 143/136	Phe 193/186	🟢	🟢	🟢
Thr 243/ Ile 236	Cys 288/ Ala 281	🟢	🟡	🟡

Table A.3. Table of analogous intramodule coevolving pairs of the transmembrane domain. Analogous coevolving pairs grouped in this table. Distance category circles – green, <5Å apart; yellow, 5-10Å apart; red, >10Å apart.

	Interdomain Interactions		Distance Catagory		
	1st Predicted Residue	2nd Predicted Residue	4C9G	Model	6GCI
H2:H3	Ile 111/Phe 103	Ser 133/126			
	Ile 111/Phe 103	Leu 140/Thr 133			
	Ille 111/Phe 103	Ala 136/129			
	Phe 107/99	Leu 140/Thr 133			
	Phe 115/107	Ala 136/129			
	Phe 107/99	Ala 136/129			
H4:H5	Leu 214/Ile 207	Val 236/Cys 229			
	Leu 214/Ile 207	Leu 233/226			
	Leu 214/Ile 207	Gly 240/233			
H6:H1	Met 29/21	Leu 308/301			
	Val 32/24	Leu 308/301			

Table A.4. Transmembrane Intermodule Highly Coevolving Interactions.

Listed are the pairs identified as highly coevolving that occur in the transmembrane region of the channel between two modules. Red circles indicate residues in the top 5% of coevolving residues and orange circles indicate top 10% and match the residues colored in Figure 2.16. Distance category circles – green, <5Å apart; yellow, 5-10Å apart; red, >10Å apart.

h12		Distance Catagory		
1st Predicted Residue	2nd Predicted Residue	4C9G	Model	6GCI
Lys 38/30	Ser 42/Ala 34			
Ser 42/Ala 34	Arg 46/38			
Ser 42/Ala 34	Leu 49/41			
Ser 42/Ala 34	Trp 87/79			
Ser 42/Ala 34	Asn 90/82			
Ile 44/36	Phe 74/66			
Arg 46/38	Leu 50/42			
Val 47/Ile 39	Phe 74/66			
Val 47/Ile 39	Pe 86/Leu 78			
Leu 50/42	Thr 77/69			
Leu 50/42	Glu 81/73			
Gln 54/46	Tyr 67/59			
Lys 66/ Arg 58	Arg 76/68			
Tyr 67/59	Asp 72/64			
Tyr 67/59	Cys 73/65			
Tyr 67/59	Arg 76/ 68			
Ala 68/Asn 60	Asp 72/64			
Ala 68/Asn 60	Arg 76/68			
Gly 69/61	Asp 72/64			
Ile 70/62	Phe 74/66			
Leu 71/Ile63	Phe 74/66			
Leu 71/Ile63	Ala 78/Thr 70			
Asp 72/64	Lys 75/Arg 67			
Asp 72/64	Arg 76/68			
Asp 72/64	Thr 79/Ala 71			
Phe 74/66	Ala 78/Thr70			
Phe 74/66	Phe 86/ Leu78			
Lys 75/Arg 67	Ala 78/Thr 70			
Arg 76/68	Thr 79/Ala 71			
Arg 76/68	Gln 80/Asp 72			
Ala 78/Thr 70	Val 83/Leu 75			
Ala 78/Thr 70	Ile 84/Met 76			
Ala 78/Thr 70	Ser 85/Ala 77			
Ala 78/Thr 70	Phe 86/Leu 78			
Val 83/Leu75	Phe 86/Leu 78			

Table A.5. Highly Coevolving Interactions of the matrix helix h12. Listed are the pairs identified as highly coevolving that occur in matrix helix h12. Red circles indicate residues in the top 5% of coevolving residues and orange circles indicate top 10% and match the residues colored in Figure 2.16. Distance category circles – green, <5Å apart; yellow, 5-10Å apart; red, >10Å apart.

h34		Distance Catagory		
1st Predicted Residue	2nd Predicted Residue	4C9G	Model	6GCI
Tyr 150/143	Thr 180/173			
Ala 151/144	Leu 189/182			
Asp 158/151	Phe 170/163			
Asp 158/151	Asp 175/168			
Gln 169/162	Lys 179/172			
Phe 170/163	Asp 175/168			
Phe 170/163	Val 176/169			
Phe 170/163	Lys 179/172			
Asn 171/164	Asp 175/168			
Asn 171/164	Lys 178/Arg 171			
Asn 171/164	Lys 179/172			
Gly 172/165	Asp 175/168			
Ile 174/167	Lys 178/ Arg 171			
Asp 175/168	Lys 178/ ARg 171			
Asp 175/168	Lys 179/172			
Val 176/169	Lys 179/172			
Val 176/169	Thr 180/173			
Tyr 177/170	Leu 181/174			
Tyr 177/170	Leu 189/182			
Lys 178/Arg 171	Leu181/174			
Lys 178/Arg 171	Lys 182/Ala 175			
Lys 179/Lys 172	Lys 182/Ala 175			
Thr 180/173	Asp 184/177			
Thr 180/173	Gly 188/181			
Leu 181/174	Asp 184/177			
Leu 181/174	Val 186/Ile 179			
Leu 181/174	Ala 187/180			
Leu 181/174	Leu 189/182			
Val 186/Ile 179	Leu 189/182			

Table A.6. Highly Coevolving Interactions of the matrix helix h34. Listed are the pairs identified as highly coevolving that occur in matrix helix h34. Red circles indicate residues in the top 5% of coevolving residues and orange circles indicate top 10% and match the residues colored in Figure 2.16. Distance category circles – green, <5Å apart; yellow, 5-10Å apart; red, >10Å apart.

h56		Distance Catagory		
1st Predicted Residue	2nd Predicted Residue	4C9G	Model	6GCI
Val 251/244	Ala 268/Ser 261			
Val 251/244	Cys 271/ Ala 264			
Val 251/244	Leu 272/Phe 265			
Val 251/244	Val 276/Ile 269			
Val 251/244	Leu 284/ 277			
Arg 254/247	Glu 279/272			
Arg 254/247	Ser 283/276			
Lys 264/257	Lys 274/Gln 267			
Tyr 265/258	Asp 270/263			
Tyr 265/258	Cys 271/Ala 264			
Tyr 265/258	Lys 274/Gln 267			
Tyr 265/258	Glu 279/272			
Asp 266/Lys 259	Lys 270/263			
Asp 266/Lys 259	Arg 273/266			
Asp 266/Lys 259	Lys 274/Gln 267			
Gly 267/Ser 260	Asp 270/ 263			
Gly 267/Ser 260	Lys 274/Gln 267			
Ala 268/Ser 261	Leu 272/Phe 265			
Phe 269/Ile 262	Leu 272/Phe 265			
Phe 269/Ile 262	Arg 273/266			
Asp 270/263	Arg 273/266			
Asp 270/263	Lys 274/Gln 267			
Asp 270/263	Ala 278/Lys 271			
Cys 271/Ala 264	Lys 274//Gln 267			
Leu 272/Phe 265	Ile 275/268			
Leu 272/Phe 265	Val 276/Ile 269			
Leu 272/Phe 265	Leu 284/277			
Arg 273/266	Ala 277/Ala 270			
Arg 273/266	Val 276/269			
Lys 274/Gln 267	Ala 277/Ala 270			
Lys 274/Gln 267	Ala 278/Lys 271			
Ile 275/268	Glu 279/272			
Ile 275/268	Ser 283/ 276			
Val 276/Ile 269	Glu 279/272			
Val 276/Ile 269	Val 281/ 274			
Val 276/Ile 269	Gly 282/ Lys 275			
Val 276/Ile 269	Ser 283/276			
Val 276/Ile 269	Leu 284/ 277			
Glu 279/272	Ser 283/276			
Glu 279/272	Lys 286/279			
Val 281/274	Leu 284/277			
Ser 283/276	Lys 286/279			

Table A.7. Highly Coevolving Interactions of the matrix helix h56. Listed are the pairs identified as highly coevolving that occur in matrix helix h56. Red circles indicate residues in the top 5% of coevolving residues and orange circles indicate top 10% and match the residues colored in Figure 2.16. Distance category circles – green, <5Å apart; yellow, 5-10Å apart; red, >10Å apart.

Appendix B – The TIM23 and TOM Complex Subunit Sequence Database Summary

Single Subunit Database			
	Fungi	**Metazoa**	**Total**
Mgr2	80	12	92
Pam16	324	260	584
Pam17	335	-	335
Pam18	161	267	428
Tim17	375	-	755
Tim17A	-	258	-
Tim17B	-	122	-
Tim21	331	332	663
Tim23	368	351	719
Tim44	244	265	509
Tim50	508	142	650
Tom20	270	319	589
Tom22	318	307	625

Table B.1. Summary of TIM23 Subunit Database. Number of individual sequences for subunits from the TIM23 or TOM Complexes.

Presequence Database	
	Total Number of Sequences
Aconitase	375
Aldehyde Dehydrogenase	156
Citrate Synthase	337
Cytochrome c Oxidase 4	309
Cytochrome c Oxidase 5a	306
Cytochrome c Oxidase 5b	299
Cytochrome c Oxidase 6a	540
Cytochrome c Oxidase 6c	187
Cytochrome c Oxidase 7c	161
Cytochrome c Oxidsae 8a	85
Dihydrolipoamide Dehydrogenase	432
Fumarate Hydratase	384
Heat Shock Protein 60	529
Isocitrate Dehydrogenase	106
Mitochondrial Ribosomal Protein, Large Subunit	103
Suppressor of Cytochrome Oxidase Deficiency	309
Succinate Dehydrogenase 1	110
Succinate Dehydrogenase 1	133
Succinate Dehydrogenase 1	413
Succinate Dehydrogenase 1	265

Table B.2. Summary of Presequence Database. Number of individual sequences for the listed proteins.

Concatonated Subunit Database		
	Fungi	**Metazoa**
Tim17 - Tim44	189	-
Tim17A - Tim44	-	125
Tim17B - Tim44	-	68
Tim23 - Mgr2	47	-
Tim23 - Pam16	163	-
Tim23 - Pam17	197	-
Tim23 - Pam18	79	-
Tim23 - Tim17	344	84
Tim23 - Tim21	155	-
Tim23 - Tim44	164	123
Tim23 - Tim50	145	-
Tim23 - Tom20	64	-
Tim23 - Tom22	187	-
Tim50 - Tom20	44	104
Tim50 - Tom22	56	-

Table B.3. Summary of Concatenated Subunit Database. Number of sequences that have been matched by species and concatenated together.

References

1. Schatz G. (1995) Mitochondria: beyond oxidative phosphorylation. *Biochimica et Biophysical Acta* **1271**, 123-126.
2. Lill R., Kispal G. (2000) Maturation of cellular Fe-S proteins: an essential function of mitochondria. *Trends in Biochemical Sciences* **25**, 352-356.
3. Newmeyer DD., Ferguson-Miller S. (2003) Mitochondria: Releasing Power for Life and Unleashing the Machineries of Death. *Cell* **112**, 481-490.
4. Lane N. and Martin W. (2010) The energetics of genome complexity. *Nature* **467**, 929-934.
5. Perkins G. and Renken C. (1997) Electron Tomography of Neuronal Mitochondria: Three-Dimensional Structure and Organization of Cristae and Membrane Contact. *Journal of Structural Biology* **119**, 260-272.
6. Hackenbrock CR. (1968) Chemical and Physical Fixation of Isolated Mitochondria in Low-Energy and High-Energy States. *Proc Natl Acad Sci USA*. **61**, 598-605.
7. Alder N.N. (2011) Biogenesis of Lipids and Proteins within Mitochondrial Membranes, in Yeagle P. (Ed.) The Structure of Biological Membranes, *CRC Press*, Boca Raton. pp. 315-377.
8. Distler AM, Kerner J., Hoppel CL (2008) Proteomics of mitochondrial inner and outer membranes. *Proteomics* **8**, 4066-4082.
9. Sickmann A., Reinders J., Wagner Y., Joppich C., Zahedi R., Meyer H.E., Schönfisch B., Perchil I., Chacinska A., Guiard B., Rehling P., Pfanner N., Meisinger C. (2003) The proteome of Saccharomyces cerevisiae mitochondria. *Proc Natl Acad Sci USA* **100**,13207-13212.
10. Lopez MF., Kristal BS., Chemokaiskaya E., Lazarev A., Shestopalov A.I., Bogdanova A., Robinson M. (2000) High-throughput profiling of the mitochondrial proteome using affinity fractionation and automation. *Electrophoresis* **21**, 3427-3440.
11. Desai N., Brown A., Amunts A., Ramakrishnan V. (2017) The structure of the yeast mitochondrial ribosome. *Science* **355**, 528-531.
12. Chacinska A., Koehler CM., Milenkovic D., Lithgow T., and Pfanner N. (2009) Importing Mitochondrial Proteins: Machineries and Mechanisms. *Cell **138***, 628-644.
13. Meisinger C., Sickmann A., Pfanner N. (2008) The Mitochondrial Proteome: From Inventory to Function. *Cell* **134**, 22-24.
14. Harayama, T., Riezman H. (2018) Understanding the diversity of membrane lipid composition. *Nature Reviews* **19**, 281-296.
15. Mårtensson CU., Doan KN., Becker T. (2017) Effects of lipids on mitochondrial functions. *Biochimica et Biophysica Acta* **1862**, 102-113.
16. Lee AG. (2011) Lipid-protein interactions. *Biochemistry Society Transactions* **39**, 761-766.
17. Palsdottir H., Lojero CG., Trumpower BL., Hunte C. (2003) Structure of the Yeast Cytochrome bc_1 Complex with a Hydroxyquinone Anion Q_o Site Inhibitor Bound. *The Journal of Biological Chemistry* **278**, 31303-31311.
18. Becker T., Horvath SE., Bottinger L., Gebert N., Daum G., Pfanner N. (2013) Role of Phosphatidylethanolamine in the Biogenesis of Mitochondrial Outer Membrane Proteins. *The Journal of Biological Chemistry* **288**, 16451-16459.
19. Schlame M., Rua D., Greenberg ML. (2000) The biosynthesis and functional role of cardiolipin. *Progress in Lipid Research* **39**, 257-288.

20. Jiang F., Ryan MT., Schlame M., Zhao M., Gu Z., Klingenberg M., Pfanner N., Greenberg ML. (2000) Absence of Cardiolipin in the crd1 Null Mutant Results in Decreased Mitochondrial Membrane Potential and Reduced Mitochondrial Function. *The Journal of Biological Chemistry* **275**, 223897-22394.
21. Krämer, R, Klingenberg, M. (1980) Enhancement of Reconstituted ADP,ATP Exchange Activity by Phosphatidylethanolamine and by Anionic Phospholipids. *FEBS Letters* **119**, 257-260.
22. Robinson NC. (1993) Functional binding of cardiolipin to cytochrome c oxidase. *Journal of Bioenergetics and Biomembranes* **25**, 153-163.
23. Arias-Cartin R., Grimaldi S., Arnoux P., Guigliarelli B., Magalon A. (2012) Cardiolipin binding in bacterial respiratory complexes: Structural and functional implications. *Biochimica et Biophysica Acta* **1817**, 1937-1949.
24. Chaban Y., Boekema EJ., Dudkina NV. (2014) Structures of mitochondrial oxidative phosphorylation supercomplexes and mechanisms for their stabilization. *Biochimica et Biophysica Acta* **1837**, 418-426.
25. Mileykovskaya E., Dowhan W. (2014) Cardiolipin-dependent formation of mitochondrial respiratory supercomplexes. *Chemistry and Physics of Lipids* **179**, 42-48.
26. Sathappa M., Alder NN. (2016) The ionization properties of cardiolipin and its variants in model bilayers. *Biochimica et Biophysica Acta* **1858**, 1362-1372.
27. van den Brink-van der Laan E., Killian A., de Kruijff B. (2004) Nonbilayer lipids affect peripheral and integral membrane proteins via changes in the lateral pressure profile. *Biochimica et Biophysica Acta* **1666**, 275-288.
28. Renner LD., Weibel DB. (2011) Cardiolipin microdomains localize to negatively curved regions of *Escherichia coli* membranes. *Proc Natl Acad Sci USA* **108**, 6264-6269.
29. Schlame M., Ren M. (2009) The role of cardiolipin in the structural organization of mitochondrial membranes. *Biochimica et Biophysica Acta* **1788**, 2080-2083.
30. Kagan VE., Chu CT., Tyurina YY., Cheikhi A., Bayir H. (2014) Cardiolipin asymmetry, oxidation and signaling. *Chem Phys Lipids* **179**, 64-69.
31. Kagan, VE., Tyurin VA., Jiang J., Tyruina Y., Ritov VB., Amoscata AA., Osipov AN., Belikova NA., Kapralov AA., Kini V., Vlasova II., Zhao Q., Zou M., Di P., Svistunenko DA., Kurnikov IV., Borisenko GG. (2005) Cytochrome c acts as a cardiolipin oxygenase required for release of proapoptotic factors. *Nature Chemical Biology* **1**, 223-232.
32. Shamas-Din A., Bindner S., Chi X., Leber B., Andrews DW., Fradin C. (2015) Distinct lipid effects on tBid and Bim activation of membrane permeabilization by pro-apoptotic Bax. *Biochemical Journal* **467**, 495-505.
33. Schlattner U., Tokarska-Schlattner M., Ramirez S., Brückner A., Kay L., Polge C., Epand RF., Lee RM., Lacombe ML., Epand RM. (2009) *Biochimica et Biophysica Acta* **1788**, 2032-2047.
34. Connerth M., Tatsuta T., Haag M., Klecker T., Westermann B., Langer T. (2012) Intramitochondrial Transport of Phosphatidic Acid in Yeast by a Lipid Transfer Protein. *Science* **338**, 815-818.
35. Tamura Y., Harada Y., Nishikawa S-I., Yamano K., Kamiya M., Shiota T., Kuroda T., Kuge O., Sesaki H., Imai K., Tomii K., Endo T. (2013) Tam41 is a CDP-Diacylglycerol Synthase Required for Cardiolipin Biosynthesis in Mitochondria. *Cell Metabolism* **17**, 709-718.

36. Kutik S., Rissler M., Guan XL., Guiard B., Shui G., Gerbert N., Heacock PN., Rehling P., Dowhan W., Wenk MR., Pfanner N., Wiedemann N. (2008) The translocator maintenance protein Tam41 is required for mitochondrial cardiolipin biosynthesis. *The Journal of Cell Biology* **183**, 1213-1221.
37. Tamura Y., Harada Y., Yamano K., Waranabe K., Ishikawa D., Ohshima C., Nishikawa SI., Yamamoto H., Endo T. (2006) Identification of Tam41 maintaining integrity of the TIM23 protein translocator complex in mitochondria. *The Journal of Cell Biology* **174**, 631-637.
38. He Q. and Greenberg ML. (2004) Post-translational regulation of phosphatidylglycerolphosphate synthase in response to inositol. *Molecular Microbiology* **53**, 1243-1249.
39. Osman C., Haag M., Wieland FT., Brügger B., Langer T. (2010) A mitochondrial phosphatase required for cardiolipin biosynthesis: the PGP phosphatase Gep4. *The EMBO Journal* **29**, 1976-1987.
40. Chang SC., Heacock PN., Mileykovskaya E., Voelker DR., Dowhan W. (1998) Isolation and Characterization of the Gene (CLS1) Encoding Cardiolipin Synthase in Saccharomyces cerevisiae. *The Journal of Biological Chemistry* **273**, 14933-14941.
41. Tuller G., Hrastnik C., Achleitner G., Schiefthaler U., Klein F., Daum G. (1998) YDL142c encodes cardiolipin synthase (Cls1p) and is non-essential for aerobic growth of Saccharomyces cerevisiae. *FEBS Letters* **421**, 15-18.
42. Beranek A., Rechberger G., Knauer H.,Wolinski H., Kohlwein SD., Leber R. (2009) Identification of a Cardiolipin-specific Phospholipase Encoded by the Gene CLD1 (YGR110W) in Yeast. *The Journal of Biological Chemistry* **284**, 11572-11578.
43. Baile MG., Sathappa M., Lu YW., Pryce E., Whited K., McCaffery JM., Han X., Alder NN., Claypool SM. (2014) Unremodeled and Remodeled Cardiolipin Are Functionally Indistinguishable in Yeast. *The Journal of Biological Chemistry* **289**, 1768-1778.
44. Xu Y., Malhotra A., Ren M., Schlame M. (2006) The Enzymatic Function of Tafazzin. *The Journal of Biological Chemistry* **281**, 39217-39224.
45. Zhong Q., Gohil VM., Ma L., Greenberg ML. (2004) Absence of Cardiolipin Results in Temperature Sensitivity, Respiratory Defects, and Mitochondrial DNA Instability Independent of pet56. *The Journal of Biological Chemistry* **279**, 32294-32300.
46. Ye C., Lou W., Li Y., Chatzispyrou IA., Huttemann M., Lee I., Houtkooper RH., Vaz FM., Chen S., Greenberg ML. (2014) Deletion of the Cardiolipin-specific Phospholipase Cld1 Rescues Growth and Life Span Defects in the Tafazzin Mutant, Implications for Barth Syndrome. *The Journal of Biological Chemistry* **289**, 3114-3125.
47. Schlame M. and Ren M. (2006) Barth syndrome, a human disorder of cardiolipin metabolism. *FEBS Letters* **580**, 5450-5455.
48. Gadir N., Haim-Vilmovsky L., Kraut-Cohen J., Gerst JE. (2011) Localization of mRNAs coding for mitochondrial proteins in the yeast *Saccharomyces cerevisiae. RNA* **17**, 1551-1565.
49. Sirrenberg C., Bauer M.F., Guiard B., Neupert W., Brunner M. (1996) Import of carrier proteins into the mitochondrial inner membrane mediated by Tim22. *Nature* **384**, 582-585.
50. Koehler CM., Jarosch E., Tokatlidis K., Schmid K., Schweyen RJ., Schatz G. (1998) Import of Mitochondrial Carriers Mediated by Essential Proteins of the Intermembrane Space. *Science* **279**, 369-373.

51. Wiedemann N., Kozjak V., Chacinska A., Schönfisch B., Rospert S., Ryan MT., Pfanner N., Meisinger C. (2003) Machinery for proteins sorting and assembly in the outer membrane. *Nature* **424**, 565-571.
52. Chacinska A., Pfannschmidt S., Wiedemann N., Kozjak V., Sanjuan Szklarz LK., Schulze-Specking A., Truscott KN., Guiard B., Meisinger C., Pfanner N. (2004) Essential role of Mia40 in import and assembly of mitochondrial intermembrane space proteins. *The EMBO Journal* **23**, 3735-3746.
53. Naoé M., Ohwa Y., Ishikawa D., Ohshima C., Nishikawa SI., Yamamoto H., Endo T. *Journal of Biological Chemistry* **279**, 47815-47821.
54. Becker T., Pfannschmidt S., Guiard B., Stojanovski D., Milenkovic D., Kutik S., Pfanner N., Meisinger C., Wiedemann N. (2008) Biogenesis of the Mitochondrial TOM Complex Mim1 promotes insertion and assembly of signal-anchored receptors. *The Journal of Biological Chemistry* **283**, 120-127.
55. Popov-Čeleketić J., Waizenegger T., Rapaport D. (2008) Mim1 Functions in an Oligomeric Form to Facilitate the Integration of Tom20 into the Mitochondrial Outer Membrane. *Journal of Molecular Biology* **376**, 671-680.
56. Chacinska A., Lind M., Frazier AE, Dudek J., Meisinger C., Geissler A., Sickmann A., Meyer HE., Truscott KN., Guiard B., Pfanner N. (2005) Mitochondrial Presequence Translocase: Switching between TOM tethering and Motor Recruitment Involves Tim21 and Tim17. *Cell* **120**, 817-829.
57. Gaume B., Klaus C., Ungermann C., Guiard B., Neupert W., Brunner M. (1998) Unfolding of preproteins upon import into mitochondria. *The EMBO Journal* **17**, 6497-6507.
58. Vögtle FN., Wortelkamp S., Zahedi RP., Becker D., Leidhold C., Gevaert K., Kellermann J., Voos W., Sickmann A., Pfanner N., Meisinger C. (2009) Global anlaysis of the mitochondrial N-proteome identifies a processing peptidase critical for protein stability. *Cell* **139**, 428-439.
59. Ou WJ., Ito A., Okazaki H., and Omura T. (1989) Purification and Characterization of a processing protease from rat liver mitochondria. *The EMBO Journal* **8**, 2605-2612.
60. Mukhopadhyay A., Hammen P., Waltner-Law M., Weiner H. (2002) Timing and structural consideration for the processing of mitochdonrial matrix space proteins by the mitochondrial processing peptidase (MPP). *Protein Science* **11**, 1026-1035.
61. Dekker PJT., Ryan MT., Brix J., Muller H., Honlinger A., Pfanner N. (1998) Preprotein Translocase of the Outer Mitochondrial Membrane: Molecular Dissection and Assembly of the General Import Pore Complex. *Molecular and Cell Biology* **18**, 6515-6524.
62. Abe Y., Shodai T., Muto T., Mihara K., Torii H., Nishikawa S., Endo T., Kohda D. (2000) Structural Basis of Presequence Recognition by the Mitochondrial Protein Import Receptor Tom20. *Cell* **100**, 551-560.
63. Yamano K., Yatsukawa Y., Esaki M., Aiken Hobbs AE., Jensen RE., Endo T. (2008) Tom20 and Tom22 Share the Common Signal Recognition Pathway in Mitochondrial Protein Import. *The Journal of Biological Chemistry* **283**, 3799-3807.
64. Mayer A., Neupert W., Lill R. (1995) Mitochondrial Protein Import: Reversible Binding of the Presequence at the *trans* Side of the Outer Membrane Drives Partial Translocation and Unfolding. *Cell* **80**, 127-137.
65. Melin J., Schulz C., Wrobel L., Bernhard O., Chacinska A., Jahn O., Schmidt B., Rehling P. (2014) Presequence Recognition by the Tom40 Channel Contributes Structural to

Precursor Translocation into the Mitochondrial Matrix. *Molecular and Cellular Biology* **34**, 3473-3485.
66. Martin J., Mahlke K., Pfanner N. (1991) Role of an Energized Inner Membrane in Mitochondrial Protein Import. *The Journal of Biological Chemistry* **266**, 18051-18057.
67. Geissler A., Krimmer T., Bömer W., Guiard B., Rassow J., Pfanner N. (2000) Membrane potential-driven protein import into mitochondria. The sorting sequence of cytochrome b_2 modulates the Δψ-dependence of translocation of the matrix-targeting sequence. *Molecular Biology of the Cell* **11**, 3977-3991.
68. Qian X., Gebert M., Hopker J. Yan M., Li J., Wiedemann N., van der Laan M., Pfanner N., Sha B. (2011) Structural Basis for the Function of Tim50 in the Mitochondrial Presequence Translocase. *J. Mol. Biol.* **411**, 513-519.
69. Geissler A., Chacinska A., Truscott KN., Wiedemann N., Brandner K., Sickmann A., Meyer HE., Meisinger C., Pfanner N., and Rehling P. (2002) The Mitochondrial Presequence Translocase: An Essential Role of Tim50 in Directing Preproteins to the Import Channel. *Cell* **111**, 507-518.
70. Mokranjac D., Paschen SA., Kozany C., Prokisch H., Hoppins SC., Nargang FE., Neupert W., and Hell K. (2003) Tim50, a novel component of the TIM23 preprotein translocase of mitochondria. *The EMBO Journal* **22**, 816-825.
71. Yamamoto H., Esaki M., Kanamori T., Tamura Y., Nishikawa SI., Endo T. (2002) Tim50 Is a Subunit of the TIM23 Complex that Links Protein Translocation across the Outer and Inner Mitochondrial Membranes. *Cell* **111**, 519-528.
72. Dekker PJT., Martin F., Maarse A.C., Bomer U., Muller H., Fuiard B., Meijer M., Rassow J., Pfanner N. (1997) The Tim core complex defines the number of mitochondrial translocation contact sites and can hold arrested preproteins in the absence of matrix Hsp70-Tim44. *The EMBO Journal* **16**, 5408-5419.
73. Truscott KN., Kovermann P., Geissler A., Merlin A., Meijer M., Driessen AJM., Rassow J., Pfanner N., Wagner R. (2001) A presequence- and voltage-sensitive channel of the mitochondrial preprotein translocase formed by Tim23. *Nature Structural Biology* **8**, 1074-1082.
74. Donzeau M., Kaldi K., Adam A., Paschen S., Wanner G., Guiard B., Bauer MF., Neupert W., and Brunner M. (2000) Tim23 Links the Inner and Outer Mitochondrial Membranes. *Cell* **101**, 401-412.
75. Maarse AC., Blom J., Keil P., Pfanner N., Meijer M. (1994) Identification of the essential yeast protein MIM17, and integral mitochondrial inner membrane protein involved in protein import. *FEBS Letters* **349**, 215-221.
76. Albrecht R., Rehling P., Chacinska A., Brix J., Cadamuro SA., Volkmer R., Guiard B., Pfanner N., Zeth K. (2006) The Tim21 binding domain connects the preprotein translocase of both mitochondrial membranes. *EMBO reports* **7**, 1233-1238.
77. van der Laan M., Wiedemann N., Mick DU., Guiard B., Rehling P., Pfanner N. (2006) A Role for Tim21 in Membrane-Potential-Dependent Preprotein Sorting in Mitochondria. *Current Biology* **16**, 2271-2276.
78. Ieva R., Heiβwolf A.K., Gebert M., Vögtle F.N., Wollweber F., Mehnert C.S., Oeljeklaus S., Warcheid B., Meisinger C., van der Laan M., Pfanner N. (2013) Mitochondrial inner membrane protease promotes assembly of presequence translocase by removing a carboxyl-terminal targeting sequence. *Nature Communications*, doi: 10.1038/ncomms3853.

79. Gebert M., Schrempp SG., Mehnert CS., HeiBwolf AK., Oeljeklaus S., Ieva R., Bohnert M., van der Malsburg K., Wiese S., Kleinschroth T., Hunte C., Meyer HE., Haferkamp I., Guiard B., Warscheid B., Pfanner N., van der Laan M. (2012) Mgr2 promotes coupling of the mitochondrial presequence translocase to partner complexes. *JCB Report*, doi: 10.1083/jcb.201110047.
80. Ieva R., Schrempp S.G., Opaliński L., Wollweber F., Höβ P., Heiβwolf A.K., Gebert M., Zhang Y., Guiard B., Rospert S., Becker T., Chacinska A., Pfanner N., van der Laan M. (2014) Mgr2 Functions as Lateral Gatekeeper for Preprotein Sorting in the Mitochondrial Inner Membrane. *Molecular Cell* **56**, 641-652.
81. Kang PJ., Ostermann J., Shilling J., Neupert W., Craig EA., Pfanner N. (1990) Requirement for hsp70 in the mitochondrial matrix for translocation and folding of precursor proteins. *Nature* **348**, 137-143.
82. Ting SY., Yan NL., Schilke BA., Craig EA. (2017) Dual interaction of scaffold protein Tim44 of mitochondrial import motor with channel-forming translocase subunit Tim23. *eLife*, doi: 10.7554/eLife.23609.
83. Josyula R., Jin Z., Fu Z., Sha B. (2006) Crystal Structure of Yeast Mitochondrial Peripheral Membrane Protein Tim44p C-terminal Domain. *J. Mol. Biol.* **359**, 798-804.
84. Mokranjac D., Bourenkov G., Hell K., Neupert W., Groll M. (2006) Strucutre and fundtion of Tim14 and Tim16, the J and J-like components of the mitochondrial protein import motor. *The EMBO Journal* **25**, 4675-4685.
85. Li Y., Dudek J., Guiard B., Pfanner N., Rehling P., Voos W. (2004) The Preseqnece Translocase-associate Protein Import Motor of Mitochondria. *The Journal of Biological Chemistry* **279**, 38047-38054.
86. Kozany C., Mokranjac D., Sichting M., Neupert W., Hell K. (2004) The J domain-related cochaperone Tim16 is a constituent of the mitochondrial TIM23 preprotein translocase. *Nature Structural & Molecular Biology* **11**, 234-241.
87. Miao B., Davis JE., Craig EA. (1997) Mge1 Functions as a Nucleotide Release Factor for Ssc1, a Mitochondrial Hsp70 of *Saccharomyces cerevisiae*. *J. Mol. Biol.* **265**, 541-552.
88. Hutu DP., Guiard B., Chacinska A., Becker D., Pfanner N., Rehling P., van der Laan M. (2008) Mitochondrial Protein Import Motor: Differential Role of Tim44 in the Recruitment of Pam17 and J-Complex to the Presequence Translocase. *Molecular Biology of the Cell* **19**, 2642-2649.
89. Popov-Čeleketić D., Mapa K., Neupert W., Mokranjac D. (2008) Active remodeling of the TIM23 complex during translocation of preproteins into mitochondrial. *The EMBO Journal* **27**, 1469-1480.
90. Ting SY., Schilke BA., Hayashi M., Craig EA. (2014) Architecture of the TIM23 Inner Mitochondrial Translocon and Interactions with the Matrix Import Motor. *The Journal of Biological Chemistry* **289**, 28689-28696.
91. Endo T., Yamano K., Kawano S. (2011) Structural insight into the mitochondrial protein import system. *Biochimica et Biophysica Acta* **1808**, 955-970.
92. Baker MJ., Frazier AE., Gulbis JM., Ryan MT. (2007) Mitochondrial protein-import machinery: correlating structure with function. *Trends in Cell Biology* **17**, 456-464.
93. Gallas MR., Dienhart MK., Stuart RA., Long RM. (2006) Characterization of Mmp37p, a *Saccharomyces cerevisiae* Mitochondrial Matrix Protein with a role in Mitochondrial Protein Import. *Molecular Biology of the Cell* **17**, 4051-4062.

94. Potting C., Wilmes C., Engmass T., Osman C., Langer T. (2010) Regulation of mitochondrial phospholipids by Ups1/PRELI-like proteins depends on proteolysis and Mdm35. *The EMBO Journal* **29**, 2888-2898.
95. Tamura T., Endo T., Iijima M., Sesaki H. (2010) Ups1p and Ups2p antagonistically regulate cardiolipin metabolism in mitochondria. *The Journal of Cell Biology* **185**, 1029-1045.
96. Weiss C., Oppliger W., Vergeres G., Demel R., Jeno P., Horst M., de Kruijff B., Schatz G., Azem A. (1999) Domain structure and lipid interaction of recombinant yeast Tim44. *Proc Natl Acad Sci USA* **96**, 8890-8894.
97. Marom M., Safonoc R., Amram S., Avneon Y., Nachliel E., Gutman M., Zohary K., Azem A., Tsfadia Y. (2009) Interaction of the Tim44 C-Terminal Domain with Negatively Charged Phospholipids. *Biochemistry* **48**, 11185-11195.
98. Malhotra K., Modek A., Nangia S., Daman TH., Gunsel U., Robinson VL., Mokranjac D., May ER., Alder NN. (2017) Cardiolipin mediates membrane and channel interactions of the mitochondrial TIM23 protein import complex receptor Tim50. *Science Advances* **3**, e1700532.
99. Horst M., Hilfiker-Rothenfluh S., Oppliger W., Schatz G. (1995) Dynamic interaction of the protein translocation systems in the inner and outer membranes of yeast mitochondria. *The EMBO Journal* **14**, 2293-2297.
100. Schülke N., Sepuri NBV., Gordon DM., Saxena S., Dancis A., Pain D. (1999) A Multisubunit Complex of Outer and Inner Mitochondrial Membrane Protein Translocases Stabilized *in Vivo* by Translocation Intermediates. *The Journal of Biological Chemistry* **274**, 22847-22854.
101. Chacinska A., Rehling P., Guiard B., Frazier AE., Schulze-Specking A., Pfanner N., Voos W., Meisinger C. (2003) Mitochondrial translocation contact sites: separation of dynamic and stabilizing elements in formation of a TOM-TIM-preprotein supercomplex. *The EMBO Journal* **22**, 5370-5381.
102. Bajaj R., Munari F., Becker S., Zweckstetter M. (2014) Interaction of the Intermembrane Space Domain of Tim23 Protein with Mitochondrial Membranes. *The Journal of Biological Chemistry* **289**, 34620-34626.
103. Sokol AM., Sztolsztener ME., Wasilewski M., Heinz E., Chacinska A. (2014) Mitochondrial protein translocases for survival and wellbeing. *FEBS Letters* **588**, 2484-2495.
104. Young JC., Hoogenraad NJ., Hartl FU. (2003) Molecular Chaperones Hsp90 and Hsp70 Deliver Preproteins to the Mitochondrial Import Receptor Tom70. *Cell* **112**, 41-50.
105. Maio RD., Barrett PJ., Hoffman EK., Barrett CW., Zharikov A., Borah A., Hu X., McCoy J., Chu CT., Burton EA., Hastings TG., Greenamyre JT. (2016) α-Synuclein binds to TOM20 and inhibits mitochondrial protein import in Parkinson's disease. *Science Translational Medicine* **8**, 342ra78.
106. Jin SM., Lazarou M., Wang C., Kane LA., Narendra DP., Youle RJ. (2010) Mitochondrial membrane potential regulates PINK1 import and proteolytic destabilization by PARL. *The Journal of Cell Biology* **191**, 933-942.
107. Lazarou M., Jin SM., Kane LA., Youle RJ. (2012) Role of PINK1 Binding to the TOM Complex and Alternate Intracellular Membranes in Recruitment and Activation of the E3 Ligase Parkin. *Developmental Cell* **22**, 320-333.

108. Bauer MF., Gempel K., Reichert AS., Rappold GA., Lichtner P., Gerbitz KD., Neupert W., Brunner M., Hofmann S. (1999) Genetic and Structural Characterization of the Mitochondrial Inner Membrane Translocase. *J. Mol. Biol.* **289**, 69-82.
109. Guo Y., Cheong NE., Zhang ZJ., De Rose R., Deng Y., Farber SA., Fernandes-Alnemri T., Alnemri ES. (2004) Tim50, a Component of the Mitochondrial Translocator, Regulates Mitochondrial Integrity and Cell Death. *The Journal of Biological Chemistry* **279**, 24813-24825.
110. Sinha D., Srivastava S., Krishna L., D'Silva P. (2014) Unraveling the intricate organization of mammalian mitochondrial presearnce translocases: existence of multiple translocases for maintenance of mitochondrial function. *Molecular and Cell Biology* **34**, 1757-1775.
111. Emtage JLT. and Jensen RE. (1993) MAS6 Encodes an Essential Inner Membrane Component of the Yeast Mitochondrial Protein Import Pathway. *The Journal of Cell Biology* **122**, 1003-1012.
112. Ahting U., Floss T., Uez N., Schneider-Lohmar I., Becker L., Kling E., Iuso A., Bender A., Hrabe de Angelis M., Gailus-Durner V., Meitinger T., Wurst W., Prokisch H., Klopstock T. (2009) Neurological phenotype and reduced lifespan in heterozygous Tim23 knockout mice, the first mouse model of defective mitochondrial import. *Biochimica et Biophysica Acta* **1787**, 371-376.
113. Roesch K., Curran SP., Tranebjaerg L., Koehler CM. (2002) Human deafness dystonia syndrome is caused by a defect in assembly of the DDP1/TIMM8a-TIMM13 complex. *Human Molecular Genetics* **11**, 477-486.
114. Rothbauer U., Hofmann S., Muhlenbein N., Paschen SA., Gerbitz KD., Neupert W., Brunner M., Bauer MF. (2001) Role of the deafness dystonia peptide 1 (DDP1) in import of human Tim23 into the inner membrane of mitochondria. *The Journal of Biological Chemistry* **276**, 37327-37334.
115. Gao SP., Sun HF., Jiang HL., Li LD., Hu X., Xu XE., Jin W. (2016) Loss of TIM50 suppresses proliferation and induces apoptosis in breast cancer. *Tumor Biology* **37**, 1279-1287.
116. Xu X., Qiao M., Zhang Y., Jiang Y., Wei P., Yao J., Gu B., Wang Y., Lu J., Wang Z., Tang Z., Sun Y., Wu W., Shi Q. (2010) Quantitative proteomics study of breast cancer cell lines isolated from a single patient: Discovery of TIMM17A as a marker for breast cancer. *Proteomics* **10**, 1374-1390.
117. Salhab M., Patani N., Jiang W., Mokbel K. (2012) High TIMM17A expression is associated with adverse patholical and clinical outcomes in human breast cancer. *Breast Cancer* **19**, 153-160.
118. Yaguchi T., Aida S., Kaul SC., Wadhwa R. (2007) Involvment of Mortalin in Cellular Senescence from the Perspective of its Mitochondrial Import, Chaperone, and Oxidative Stress Managemnet Functions. *Ann. N.Y. Acad. Sci.* **1100**, 306-311.
119. Craig EA., Kramer J., Kosic-Smithers J. (1987) SSC1, a member of the 70-kDa heat shock protein multigene family of *Saccharomyces cerevisiae*, is essential for growth. *Proc. Natl. Acad. Sci USA* **84**, 4156-4160.
120. Craig EA., Kramer J., Shilling J., Werner-Washburne M., Holmes S., Kosic-Smithers J., Nicolet CM. (1989) SSC1, an Essential Member of the Yeast HSP70 Multigene Family, Encodes a Mitochondrial Protein. *Molecular and Cellular Biology* **9**, 3000-3008.

121. Bonora E., Evangelisti C., Bonichon F., Tallini G., Romeo G. (2006) Novel germline variants identified in the inner mitochondrial membrane transporter TIMM44 and their role in predisposition to oncocytic thyroid carcinomas. *British Journal of Cancer* **95**, 1529-1536.
122. Wada J. and Kanwar YS. (1998) Characterization of mammalian translocase of inner mitochondrial membrane (Tim44) isolated from diabetic newborn mouse kidney. *Proc. Natl. Acad. Sci. USA* **95**, 144-149.
123. Jubinsky PT., Short MK., Mutema G., Morris RE., Ciraolo GM., Li M. (2005) Magmas expression in neoplastic human prostate. *J Mol Hist* **36**, 69-75.
124. Mehawej C., Delahodde A., Legeai-Mallet L., Delague V., Kaci N., Desvignes JP., Kibar Z., Capo-Chichi JM., Chouery E., Munnich A., Cormier-Daire V., Mégarbané A. (2014) The Impairment of MAGMAS Function in Human is Responsible for a Severe Skeletal Dysplasia. *PLoS Genetics*, 10. e1004311.
125. Lu WJ., Lee NP., Kaul SC., Lan F., Poon RTP., Wadhwa R., Luk JM. (2011) Mortalin-p53 interaction in cancer cells is stress dependent and constitutes a selective target for cancer therapy. *Cell Death and Differentiation* **18**, 1046-1056.
126. Zarsky V. and Dolezal P. (2016) Evolution of the Tim17 protein family. *Biology Direct* **11-54**, doi: 10.1186/s13062-016-0157-y
127. Mick DU., Dennerlein S., Wiese H., Reinhold R., Pacheu-Grau D., Lorenzi I., Sasarman F., Weraarpachai W., Shoubridge EZ., Warscheid B., Rehling P. (2012) MITRAC Links Mitochondrial ProteinTranslocation to Respiratory-Chain Assembly and Translational Regulation. *Cell* **151**, 1528-1541.
128. Chung YM., Kim JS., Yoo YD. (2006) A novel protein, Romo1, induces ROS production in the mitochondria. *Biochemical and Biophysical Research Communications* **347**, 649-655.
129. Chung JS., Park S., Park SH., Park ER., Cha PH., Kim BY., Chung YM., Woo SR., Han CJ., Kim BS., Suh KS., Jang JJ., Lee K., Choi DW., Lee S., Lee GY., Hahm KB., Shin JA., Kim BS., Noh KH., Kim TW., Lee KH., Yoo YD. (2012) Overexpression of Romo1 promotes production of reactive oxygen species and invasiveness of hepatic tumor cells. *Gastroenterology* **143**, 1084-1094.
130. Botstein D., Fink GR. (2011) Yeast: An Experimental Organism for 21st Century Biology. *Genetics* **189**, 695-704.
131. Kovermann P., Truscott K.N., Guiard B., Rehling P., Sepuri N.B., Müller H., Jensen R.E., Wagner R., Pfanner N. (2002) Tim22, the Essential Core of the Mitochondrial Protein Insertion Complex, Forms a Voltage-Activated and Signal-Gated Channel. *Cell* **9**, 363-373.
132. Kaldi K., Bauer MF., Sirrenberg C., Neupert M., Brunner M. (1998) Biogenesis of Tim23 and Tim17, integral components of the TIM machinery for matrix-targeted preproteins. *The EMBO Journal* **17**, 1569-1576.
133. Davis AJ., Sepuri NB., Holder J., Johnson AE., Jensen RE. (2000) Two Intermembrane Space TIM Complexes Interact with Difference Domains of Tim23p during Its Import into Mitochondria. *The Journal of Cell Biology* **150**, 1271-1282.
134. Bauer MF., Hofmann S., Neupert W., Brunner M. (2000) Protein translocation into mitochondria: the role of TIM complexes. *Trends in Cell Biology* **10**, 25-31.

135. Curran SP., Leuenberger D., Schmidt E., Koehler CM. (2002) The role of the Tim8p-Tim13p complex in a conserved import pathway for mitochondrial polytopic inner membrane proteins. *The Journal of Cell Biology* **158**, doi: 10.1083/jcb.200205124.
136. Rehling P., Model K., Brandner K., Kovermann P., Sickmann A., Meyer HE., Kühlbrandt W.,Wagner R., Truscott KN., Pfanner N. (2003) Protein Insertion into the Mitochondrial Inner Membrane by a Twin-Pore Translocase. *Science* **299**, 1747-1751.
137. Gebert N., Chacinska A., Wagner K., Guiard B., Koehler CM., Rehling P., Pfanner N., Wiedemann N. (2008) Assemble of the three small Tim proteins precedes docking to the mitochondrial carrier translocase. *EMBO reports* **9**, 548-554.
138. Wagner K., Gebert N., Guiard B., Brandner K. Truscott KN., Wiedemann N., Pfanner N., Rehling P. (2008) Assembly Pathway of the Mitochondrial Carrier Translocase Involves Four Preprotein Translocases. *Molecular and Cellular Biology* **28**, 4251-4260.
139. Kurz M., Martin H., Rassow J., Pfanner N., Ryan MT. (1999) Biogenesis of Tim Proteins of the Mitochondrial Carrier Import Pathway: Differential Targeting Mechanisms and Crossing Over with the Main Import Pathway. *Molecular Biology of the Cell* **10**, 2461-2474.
140. Gebert N., Gebert M., Oeljeklaus S., von der Malsburg K., Stroud DA., Kulawiak B., Wirth C., Zahedi RP., Dolezal P., Wiese S., Simon O., Schulze-Specking A., Truscott KN., Sickmann A., Rehling P., Guiard B., Hunte C., Warscheid B., van der Laan M., Pfanner N., Wiedemann N. (2011) Dual Function of Sdh3 in the Respiratory Chain and TIM22 Protein Translocase of the Mitochondrial Inner Membrane. *Molecular Cell* **44**, 811-818.
141. Chacinska A., Koehler CM., Milenkovic D., Lithgow T., Pfanner N. (2009) Importing Mitochondrial Proteins: Machineries and Mechanisms. *Cell* **138**, 628-644.
142. Schmidt O., Pfanner N., Meisinger C. (2010) Mitochondrial protein import: from proteomics to functional mechanism. *Nature Reviews: Molecular Cell Biology* **11**, 655-667.
143. Hines V., Brandt A., Griffiths G., Horstmann H., Brutsch H., Schatz G. (1990) Protein import into yeast mitochondria is accelerated by the outer membrane protein MAS70. *The EMBO Journal* **9**, 3191-3200.
144. Webb CT., Gorman MA., Lazarou M., Ryan MT., Gulbis JM. (2006) Crystal Structure of the Mitochondrial Chaperone Tim9-10 Reveals a Six-Bladed α-Propeller. *Molecular Cell* **21**, 123-133.
145. Davis AJ., Alder NN., Jensen RE., Johnson AE. (2007) The Tim9p/Tim10p and Tim8p/13p Complexes Bind to Specific Sites on Tim23p during Mitochondrial Protein Import. *Molecular Biology of the Cell* **18**, 475-486.
146. Truscott KN., Wiedemann N., Rehling P., Muller H., Meisinger C., Pfanner N., Guiard B. (2002) Mitochondrial Import of the ADP/ATP Carrier: the Essential TIM Complex of the Intermembrane Space Is Required for Precursor Release from the TOM Complex. *Molecular and Cellular Biology* **22**, 7780-7789.
147. Palmieri F. (2013) The mitochondrial transporter family (SLC25): identification, properties, and physiopathology. *Molecular Aspects of Medicine* **34**, 465-484.
148. Saraste M., Walker JE. (1982) Internal sequence repeats and the path of polypeptide in mitochondrial ADP/ATP translocase. *FEBS Letters* **144**, 250-254.
149. Klingenberg M. (2008) The ADP and ATP transport in mitochondria and its carrier. *Biochimica et Biophysical Acta* **1778**, 1978-2021.

150. Pfaff E., Heldt HW., Klingenberg M. (1969) Adenine Nucleotide Translocation of Mitochondria: Kinetics of the Adenine Nucleotide Exchange. *European Journal of Biochemistry* **10**, 484-493.
151. Weidemann MJ., Erdelt H., Klingenberg M. (1970) Adenine Nucleotide Translocation of Mitochondria: Identification of Carrier Sites. *European Journal of Biochemistry* **16**, 313-335.
152. Erdelt H., Weidemann MJ., Buchholz M., Klingenberg M. (1972) Some Principle Effects of Bongkrekic Acid on the Binding of Adenine Nucleotides to Mitochondrial Membranes. *European Journal of Biochemistry* **30**, 107-122.
153. Klingenberg M., Buchholz M. (1973) On the Mechanism of Bongkrekate Effect on the Mitochondrial Adenine-Nucleotide Carrier as Studied through the Binding of ADP. *European Journal of Biochemistry* **38**, 346-358.
154. Buchanan BB., Eiermann W., Riccio P., Aquila H., Klingenberg M. (1976) Antibody evidence for different conformational states of ADP,ATP translocator protein isolated from mitochondria. *Proc. Natl. Acad. Sci. USA* **73**, 2280-2284.
155. Pebay-Peyroula E., Dahout-Gonzalez C., Kahn R., Trézéguet V., Lauquin GJ., Brandolin G. (2003) Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractyloside. *Nature* **426**, 39-44.
156. Nury H., Dahout-Gonzalez C., Trézéguet V., Lauguin G., Brandolin G., Pebay-Peyroula E. (2005) Structural basis for lipid-mediated interactions between mitochondrial ADP/ATP carrier monomers. *FEBS Lett*. **579**, 6031-6036.
157. Ruprecht JJ., Hellawell AM., Harding M., Crichton PG., McCoy AJ., Kunji ERS. (2014) Structures of Yeast Mitochondrial ADP/ATP Carriers Support a Domain-Based Alternating-Access Transport Mechanism. *Proc Natl Acad Sci USA*. **111**, E426-E434.
158. Kunji ER, Harding M. (2003) Projection Structure of the Atractyloside-Inhibited Mitochondrial ADP/ATP Carrier of Saccharomyces Cerevisiae*. J. Biol. Chem.* **39**, 36985-36988.
159. Ruprecht JJ., King MS., Zögg T., Aleksandrova AA., Pardon E., Crichton PG., Steyaert J., Kunji ERS. (2019) The Molecular Mechanism of Transport by the Mitochondrial ADP/ATP Carrier. *Cell* **176**, 435-447.
160. Vignais PV, Vignais PM, Defaye G. (1973) Adenosine Diphosphate Translocation in Mitochondria. Nature of the Receptor Site for Carboxyatractyloside (Gummiferin). *Biochemistry* **12**, 1508-1519.
161. Henderson PJ., Lardy HA. (1970) Bongkrekic Acid. An inhibitor of the adenine nucleotide translocator of mitochondria. *J Biol Chem*. **245**, 1319-1326.
162. Nelson DR, Felix CM, Swanson JM. (1998) Highly Conserved Charge-pair Networks in the Mitochondrial Carrier Family. *J Mol Biol.* **277**, 285-308.
163. Robinson AJ., Overy C., Kunji ERS. (2008) The mechanism of transport by mitochondrial carriers based on analysis of symmetry. *Proc Natl Acad Sci USA* **105**, 17766-17771.
164. King MS., Kerr M., Crichton PG., Springett R., Kunji ERS. (2016) Formation of a cytoplasmic salt bridge network in the matrix state is a fundamental step in the transport mechanism of the mitochondrial ADP/ATP carrier. *Biochimica et Biophysica Acta* **1857**, 14-22.
165. Jardetsky O. (1966) Simple allosteric model for membrane pumps. *Nature* **211**, 969-970.

166. Bamber L., Harding M., Monne M., Slotboon, D-J., Kunji, ERS. (2007) The yeast mitochondrial ADP/ATP carrier functions as a monomer in mitochondrial membranes. *Proc Natl Acad Sci USA* **104**, 10830-10834.
167. Kunji ERS., and Crichton PG. (2010) Mitochondrial carriers function as monomers. *Biochimica et Biophysica Acta* **1797**, 817-831.
168. Robinson AJ., Kunji, ERS. (2006) Mitochondrial carriers in the cytoplasmic state have a common substrate binding site. *Proc Natl Acad Sci USA* **103**, 2617-2622.
169. Dalbon P., Brandolin G., Boulay F., Hoppe J., Vignais PV. (1988) Mapping of the nucleotide-binding sites in the ADP/ATP carrier of beef heart mitochondria by photolabeling with 2-azido[alpha-32P]adenosine diphosphate. *Biochemistry* **12**, 5141-5149.
170. Mayinger P., Winkler E., Klingenberg MK. (1989) The ADP/ATP carrier from yeast (AAC-2) is uniquely suited for the assignment of the binding center by photoaffinity labeling. *FEBS Letters* **244**, 421-426.
171. Monne M., Palmieri F., Kunji ERS. (2013) The substrate specificity of mitochondrial carriers: Mutagenesis revisited. *Molecular Membrane Biology* **30**, 149-159.
172. Kunji ERS., Robinson AJ. (2006) The conserved substrate binding site of mitochondrial carriers. *Biochimica et Biophysica Acta* **1757**, 1237-1248.
173. Dehez F., Pebay-Peyroula E., Chipot C. (2008) Binding of ADP in the mitochondrial ADP/ATP carrier is driven by an electrostatic funnel. *Journal of the American Chemical Society* **130**, 12725-12733.
174. Wang Y., Tajikhorshid E. (2008) Electrostatic funneling of substrate in mitochondrial inner membrane carriers. *Proc Natl Acad Sci USA* **105**, 9598-9603.
175. Rey M., Man P., Clemencon B., Trezguet V., Brandolin G., Forest E., Pelosi L. (2010) Conformational dynamics of the bovine mitochondrial ADP/ATP carrier isoform 1 revealed by hydrogen/deuterium exchange coupled to mass spectrometry. *J. Biol. Che* **285**, 34981-34990.
176. Rey M., Forest E., Pelosi L. (2012) Exploring the conformational dynamics of the bovine ADP/ATP carrier in mitochondria. *Biochemistry* **51**, 9727-9735.
177. Bruchweiler S., Yang Q., Chou JJ. (2015) Substrate-modulated ADP/ATP-transporter dynamics revealed by NMR relaxation dispersion. *Nat. Struct. Mol. Biol.* **22**, 636-641.
178. Beyer K., Klingenberg M. (1985) ADP/ATP Carrier Protein from Beef Heart Mitochondria Has High Amounts of Tightly Bound Cardiolipin, AS Revelared by ^{31}P Nuclear Magnetic Resonance. *Biochemistry* **24**, 3821-3826.
179. Hoffmann B., Stöckl A., Schlame M., Beyer K., Klingenberg M. (1994) The Reconstituted ADP/ATP Carrier Activity Has an Absolute Requirement for Cardiolipin as Shown in Cysteine Mutants. *The Journal of Biological Chemistry* **269**, 1940-1944.
180. Kunji ERS., Aleksandrova A., King MS., Majd H., Ashton VL., Cerson E., Springett R., Kibalchenko M., Tavoulari S., Crichton PG., Ruprecht JJ. (2016) The transport mechanism of the mitochondrial ADP/ATP carrier. *Biochimica et Biophysica Acta* **1863**, 2379-2393.
181. Duncan AL., Ruprecht JJ., Kunji ERS., Robinson AJ. (2018) Cardiolipin dynamics and binding to conserved residues in the mitochondrial ADP/ATP carrier. *Biochimica et Biophysica Acta – Biomembranes* **1860**, 1035-1045.

182. Tamura K., Hayashi S. (2017) Atomistic modeling of alternating access of a mitochondrial ADP/ATP membrane transporter with molecular simulations. *PLoS One* **12**, e0181489.
183. Göbel U., Sander C., Schneider R., Valencia A. (1994) Correlated mutations and residue contacts in proteins. *Proteins* **18**, 309-317.
184. Casari G., Sander C., Valencia A. (1995) A method to predict functional residues in proteins. *Nat Struct Biol.* **2**, 171-178.
185. Neher E. (1994) How frequent are correlated changes in families of protein sequences?. *Proc Natl Acad Sci USA* **91**, 98-102.
186. Fodor AA., Aldrich RW. (2004) Influence of Conservation on Calculations of Amino Acid Covariance in Multiple Sequence Alignments. *Proteins* **56**, 211-221.
187. Tress ML., Valencia A. (2010) Predicted residue-residue contacts can help the scoring of 3D models. *Proteins* **78**, 1980-1991.
188. Rigden DJ. (2002) Use of covariance analysis for the prediction of structural domain boundaries from multiple sequence alignments. *Protein Engineering, Design and Selection* **15**, 65-77.
189. Lapedes AS, Giraud BG., Liu LC., Stormo GD. (1998) A maximum entropy formalism for disentangling chains of correlated sequence positions. No. LA-UR-98-1094. Los Alamos National Lab., NM (US).
190. Thomas J., Ramakrishnan N., Bailey-Kellogg C. (2005) Graphical models of residue coupling in protein families. In: BIOKDD '05: Proceedings of the 5th International Workshop on Bioinformatics, ACM, New York, NY, 12–20.
191. Weigt M., White RA., Szurmant H., Hoch JA., Hwa T. (2009) Identification of direct residue contacts in protein–protein interaction by message passing. *Proc Nat Acad of Sci USA* **106**, 67-72.
192. Balakrishnan S., Kamisetty H., Carbonell JG., Lee SI., Langmead CJ. (2011) Learning generative models for protein fold families. *Proteins: Structure, Function, and Bioinformatics* **79**, 1061-1078.
193. Morcos F., Pagnani A., Lunt B., Bertolino A., Marks DS., Sander C., Zecchina R., Onuchic JN., Hwa T., Weigt M. (2011) Direct-coupling analysis of residue coevolution captures native contacts across many protein families. *Proc Natl Acad Sci USA* **108**, E1293–E1301.
194. Jones DT., Buchan DWA., Cozzetto D., Pontil M. (2012) PSICOV: precise structural contact prediction using sparse inverse covariance estimation on large multiple sequence alignments. *Bioinformatics* **28**, 184-190.
195. Magnus E., Lövkvist C., Lan Y., Weigt M., Aurell E. (2012) Improved contact prediction in proteins: Using pseudolikelihoods to infer Potts models. *Phys Rev E.* **87**, 012707.
196. Kamisetty H., Ovchinnikov S., Baker D. (2013) Assessing the utility of coevolution-based residue-residue contact predictions in a sequence- and structure-rich era. *Proc Natl Acad Sci USA.* **110**, 15674-15679.
197. Wang S., Sun S., Li Z., Zhang R., Xu J. (2017) Accurate De Novo Prediction of Proteins Contact Map by Ultra-Deep Learning Model. *PLoS Comput Biol.* **13**, e1005324.

198. Källberg M., Wang H., Wang S., Peng J., Wang Z., Lu H., Xu J. (2012) Template-based protein structure modeling using the RaptorX web server. *Nature Protocols* **7**, 1511-1522.
199. Ovchinnikov S., Kinch L., Park H., Liao Y., Pei J., Kim DE., Kamisetty H., Grishin NV., Baker D. (2015) Large-scale determination of previously unsolved protein structures using evolutionary information. *Elife* **4**, e09248.
200. Hopf TA., Colwell LJ., Sheridan R., Rost B., Sander C., Marks DS. (2012) Three-Dimensional Structures of Membrane Proteins from Genomic Sequencing. *Cell* **149**, 1607-1621.
201. Söding J. (2005) Protein homology detection by HMM-HMM comparison. *Bioinformatics* **21**, 951-960.
202. Remmert M., Biegert A., Hauser A., Söding. J. (2012) HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nature Methods* **9**, 173-175.
203. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. (2000) The Protein Data Bank. Nucleic Acids Research **28**, 235-242.
204. Rawi R., Kunji K., Haoudi A., Bensmail H. (2015) Coevolution Analysis of HIV-1 Envelope Glycoprotein Complex. *PLoS ONE* **10**, e0143245.
205. Crichton PG., Lee L., Ruprecht JJ., Cerson E., Thangaratnarajah C., King MS., Kunji ERS. (2015) Trends in Thermostability Provide Information on the Nature of Substrate, Inhibitor and Lipid Interactions with Mitochondrial Carriers. *The Journal of Biological Chemistry* **290**, 8206-8217.
206. Cappello AR., Curcio R., Miniero DV., Stipani T., Robinson AJ., Kunji ERS., Palmieri F. (2006) Functional and Structural Role of Amino Acid Residues in Even-numbered Transmembrane α-Helices of the Bovine Mitochondrial Oxoglutarate Carrier. *Journal of Molecular Biology* **363**, 51-62.
207. Lui Y., Engelman DM., Gerstein M. (2002) Genomic analysis of membrane protein families: abundance and conserved motifs. *Genome Biology* **3**, research0054.1-research0054.12.
208. Clare A., King RD. (2003) Predicting gene function in Saccharomyces cerevisiae. *Bioinformatics* **19**, ii42-ii49.
209. Bauer MF., Sirrenberg C., Neupert W., Brunner M. (1996) Role of tim23 as Voltage Sensor and Presequence Receptor in Protein Import into Mitochondria. *Cell* **87**, 33-41.
210. Tamura Y., Harada Y., Shiota T., Yamano K., Watanabe K., Yokota M., Yamamoto H., Sesaki H., Endo T. (2009) Tim23-Tim50 pair coordinates functions of translocators and motor proteins in mitochondrial protein import. *Journal of Cell Biology* **184**, 129-141.
211. Kubrish M., Keil P., Rassow J., Dekker PJT., Blom J., Meijer M. Pfanner N. (1994) the polytopic mitochondrial inner membrane proteins MIM17 and MIM23 operate at the same preprotein import site. *FEBS Letters* **349**, 222-228.
212. Senes A., Engel DE., DeGrado WF. (2004) Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Current Opinion in Structural Biology* **14**, 465-479.
213. Demishtein-Zohary K., Marom M., Neupert W., Mokranjac D., Azem A. (2015) GxxxG motifs hold the TIM23 complex together. *FEBS Journal* **282**, 2178-2186.

214. Alder NN., Jensen RE., Johnson AE. (2008) Fluorescence Mapping of Mitochondrial TIM23 Complex Reveals a Water-Facing, Substrate-Interacting Helix Surface. *Cell* **132**, 439-450.
215. Malhotra K., Sathappa M., Landin JS., Johnson AE., Alder NN. (2013) Structural changes in the mitochondrial Tim23 channel are coupled to the proton-motive force. *Nature Structural & Molecular Biology* **20**, 965-972.
216. Denkert N., Schendzielorz AB., Barbot M., Versemann L., Richter F., Rehling P., Meinecke M. (2017) Cation selectivity of the presequence translocase channel Tim23 is crucial for efficient protein import. *eLife*, doi: 10.7554/eLife.28324.
217. Ramesh A., Peleh V., Martinez-Caballero S., Wollweber F., Sommer F., van der Laan M., Schroda M., Alexander RT., Campo ML., Herrmann JM. (2016) A disulfide bond in the TIM23 complex is crucial for voltage gating and mitochondrial protein import. *The Journal of Cell Biology* **214**, 417-431.
218. Demishtein-Zohary K., Gunsel U., Marom M., Banerjee R., Neupert W., Azem A., Mokranjac D. (2017) Role of Tim17 in coupling the import motor to the translocation channel of the mitochondrial presequence translocase. *eLife,* 10.7554/elife.22696
219. Alder NN., Sutherland J., Buhring AI., Jensen RE., Johnson AE. (2007) Quaternary Structure of the Mitochondrial Tim23 Complex Reveals Dynamic Association between Tim23p and Other Subunits. *Molecular Biology of the Cell* **19**, 159-170.
220. Simkovic F., Ovchinnikov S., Baker D., Rigden DJ. (2017) Applications of contact predictions to structural biology. *International Union of Crystallography Journal* **4**, 291-300.
221. Benson DA., Cavanaugh M., Clark K., Karsch-Mizrachi IK., Lipman DJ., Ostell J., Sayers EW. (2013) GenBank. *Nucleic Acids Research* **41**, D36-D42.
222. The UniProt Consortium. (2108) UniProt: the universal protein knowledgebase. *Nucleic Acids Research* **45**, D158-D169.
223. Ashkenazy H., Abadi S., Martz E., Chay O., Mayrose I., Pupko T., Ben-Tal N. (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucl. Acids Res* **44**, W344-W350.
224. Celniker G., Nimrod G., Ashkenazy H., Glaser F., Martz E., Mayrose I., Pupko T., Ben-Tal N. (2013) ConSurf: Using Evolutionary Data to Raise Testable Hypotheses about Protein Function. *Isr. J. Chem*. **53**, 199-206.
225. Ashkenazy H., Erez E., Martz E., Pupko T., Ben-Tal N. (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucl. Acids Res.* **38**, W529-W533.
226. Landau M., Mayrose I., Rosenberg Y., Glaser F., Martz E., Pupko T., Ben-Tal N. (2005) ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucl. Acids Res.* **33**, W299-W302.
227. Glaser F., Pupko T., Paz I., Bell R.E., Bechor D., Martz E., Ben-Tal N. (2003) ConSurf: Identification of Functional Regions in Proteins by Surface-Mapping of Phylogenetic Information. *Bioinformatics* **19**, 163-164.
228. Mayrose I., Graur D., Ben-Tal N., Pupko T. (2004) Comparison of site-specific rate-inference methods for protein sequences: empirical Bayesian methods are superior. *Mol. Biol. Evol*. **21**, 1781-1791.
229. van der Laan M., Meinecke M., Dudek J., Hutu DP., Lind M., Perschil I., Guiard B., Wagner R., Pfanner N., Rehling P. (2007) Motor-free mitochondrial presequence

translocase drives membrane integration of preproteins. *Nature Cell Biology* **9**, 1152-1159.
230. Wittig I., Braun HP., Schägger H. (2006) Blue native PAGE. *Nature Protocols* **1**, 418-428.
231. Borch J., Hamann T. (2009) The nanodisc: a novel tool for membrane protein studies. *Biological Chemistry* **390**, 805-814.
232. Bayburt TH., Grinkova YV., Sligar SG. (2002) Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Letters* **2**, 853-856.
233. Bayburt TH., Sligar SG. (2003) Self-assembly of single integral membrane proteins into soluble nanoscale phospholipid bilayers. *Protein Science* **12**, 2476-2481.
234. Ranaghan MJ., Schwall CT., Alder NN., Brige RR. (2011) Green Proteorhodopsin Reconstituted into Nanoscale Phospholipid Bilayers (Nanodiscs) as Photoactive Monomers. *Journal of the American Chemical Society* **133**, 18318-18327.
235. Schwall CT., Greenwood VL., Alder NN. (2012) The stability and activity of respiratory complex II is cardiolipin-dependent. *Biochimica et Biophysica Acta* **1817**, 1588-1596.
236. Alami M., Dalal K., Lelj-Garolla B., Sligar SG., Duong F. (2007) Nanodiscs unravel the interaction between the SecYEG channel and its cytosolic partner SecA. *The EMBO Journal* **26**, 1995-2004.
237. Dalal K., Nguyen N., Alami M., Tan J., Moraes TF., Lee WC., Maurus R., Sligar SG., Brayer GD., Duong F. (2009) Structure, Binding, and Activity of Syd, a SecY-interaction Protein. *The Journal of Biological Chemistry* **284**, 7897-7902.
238. Denisov IG., Grinkova YV., Lazarides AA., Sligar SG. (2004) Directed Self-Assembly of Monodisperse Phospholipid Bilayer Nanodiscs with Controlled Size. *Journal of the American Chemical Society* **126**, 3477-3487.
239. Ikon N., Su B., Hsu FF., Forteand TM., Ryan RO. (2015) Exogenous cardiolipin localizes to mitochondria and prevents TAZ knockdown-induced apoptosis in myeloid progenitor cells. *Biochem Biophys Res Cummun*. **464**, 580-585.
240. Pfeiffer K., Gohil V., Stuart RA., Hunte C., Brandt U., Greenberg ML., Schägger H. (2003) Cardiolipin stabilizes respiratory chain supercomplexes. *The Journal of Biological Chemistry*, **278** 52873-52880.
241. Daum G., Böhni PC., Scharz G. (1982) Import of proteins into mitochondria. cytochrome b2 and cytochrome c peroxidase are located in the Intermembrane space of yeast mitochondria. *The Journal of Biological Chemistry* **257**, 13028-13033.
242. Brandner K., Mick DU., Frazier AE., Taylor RD., Meisinger C., Rehling P. (2005) Taz1, an outer mitochondrial membrane protein, affects stability and assembly of inner membrane protein complexes: implications for Barth Syndrome. *Molecular Biology of the Cell* **16**, 5202-5214.
243. Zhang M., Mileykovskaya E., Dowhan W. (2002) Gluing the respiratory chain together. Cardiolipin is required for supercomplex formation in the inner mitochondrial membrane. *The Journal of Biological Chemistry* **277**, 43553-43556.
244. Zhang M., Mileykovskaya E., Dowhan W. (2005) Cardiolipin in essential for organization of complexes III and IV into a supercomplex in intact yeast mitochondria. *The Journal of Biological Chemistry* **280**, 29403-29408.

245. Allison DS., Schatz G. (1986) Artificial mitochondrial presequences. *Proc. Natl. Acad. Sci USA* **83**, 9011-9015.
246. Saraste M. (1999) Oxidative Phosphorylation at the fin de siècle. *Science* **283**, 1488-1493.
247. Iverson TM., Luna-Chavez C., Cecchini G., Rees DC. (1999) Structure of the *Escherichia coli* fumarate reductase respiratory complex. *Science* **284**, 1961-1966.
248. Lancaster CRD., Kröger A., Auer M., Michel H. (1999) Structure of fumarate reductase from *Wolinella Succinogenes* at 2.2Å resolution. *Nature* **402**, 377-385.
249. Oyedotun KS., Lemire BD. (2004) Quaternary Structure of the *Saccharomyces cerevisiae* Succinate Dehydrogenase. Homology Modeling, cofactor docking, and molecular dynamics simulation studies. *The Journal of Biological Chemistry* **279**, 9424-9431.
250. Cecchini G. (2003) Function and Structure of Complex II of the Respiratory Chain. *Annu. Rev. Biochem* **72**, 77-109.
251. Maklashina E., Cecchini G. (2010) The quinone-binding and catalytic site of complex II. *Biochimica et Biophysica* Acta **1797**, 1877-1882.
252. Larson AM., Polson J., Fontana RJ., Davern TJ., Lalani E., Hynan LS., Reisch JS., Schiødt FV., Ostapowicz G., Shakil AO., Lee WM., the Acute Liver Failure Study Group. (2005) Acetaminophen-Induced Acute Liver Failure: Results of a United States Multicenter, Prospective Study. *Hepatology* **42**, 1364-1372.
253. Birge RB., Bartolone JB., Hart SG., Nishanian EV., Tyson CA., Khairallah EA., Cohen SD. (1990) Acetaminophen hepatotoxicity: correspondence of selective protein arylation in human and mouse liver in vitro, in culture and in vivo. *Toxicology and Applied Pharmacology* **105**, 472-482.
254. Jaeschke H., McGill MR., Ramachandran A. (2012) Oxidant stress, mitochondria, and cell death mechanisms in drug-induced injury: lessons learned from acetaminophen hepatotoxicity. *Drug Metabolism Reviews* **44**, 88-106.
255. Han D., Dara L., Win S., Than TA., Yuan L., Abbasi SQ., Liu ZX., Kaplowitz N. (2013) Regulation of drug-induced liver injury by signal transduction pathways: critical role of mitochondria. *Trends in Pharmacological Sciences* **34**, P243-P253.
256. Burke AS., MacMillaan-Crow LA., Hinson JA. (2010) Reactive Nitrogen Species in Acetaminophen-Induced Mitochondrial Damage and Toxicity in Mouse Hepatocytes. *Chemical Research in Toxicity* **23**, 1286-1292.
257. Dahlin DC., Miwa GT., Lu AT., Nelson SD. (1984) N-acetyl-p-benzoquinone imine: a cytochrome P-450-mediated oxidation product of acetaminophen. *Proc Natl Acad Sci USA* **81**, 1327-1331.
258. LoGuidice A., Boelsterli UA. (2011) Acetaminophen overdose-induced liver injury in mice is mediated by peroxynitrite independently of the cyclophilin D-regulated permeability transition. *Hepatology* **54**, 969-978.
259. Nelson SD., Tirmenstein SA., Rashed MS., Myers TG. (1991) Acetaminophen and protein thiol modification. *Advances in Experimental Medicine and Biology* **283**, 579-588.
260. Cover C., Mansouri A., Knight TR., Bajt ML., Lemasters JJ., Pessayre D., Jaeschke H. (2005) Peroxynitrite-induced mitochondrial and endonuclease-mediated nuclear DNA damage in acetaminophen hepatoxicity. *Journal of Pharmacology and Experimental Therapeutics* **315**, 879-887.

261. Abdelmegeed MA., Jang S., Banergee A., Hardwick JP., Song BJ. (2013) Robust protein nitration contributes to acetaminophen-induced mitochondrial dysfunction and acute liver injury. *Free Radical Biology and Medicine* 60, 211-222.
262. Ramsay RR., Rashed MS., Nelson SD. (1989) In Vitro Effects of Acetaminophen Metabolites and Analogs on the Respiration of Mouse Liver Mitochondria. *Archives of Biochemistry and Biophyiscs* **273**, 449-457.
263. Burcham PC., Harman AW. (1991) Acetaminophen Toxicity Results in Site-specific Mitochondrial Damage in Isolated Mouse Hepatocytes. *The Journal of Biological Chemistry* **266**, 5049-5054.
264. Lee KK., Fujimoto K., Zhang C., Schwall CT., Alder NN., Pinkert CA., Krueger W., Rasmussen T., Boelsterli UA. (2013) Isoniazid-induced cell death is precipitated by underlying mitochondrial complex I dysfunction in mouse hepatocytes. *Free Radical Biology and Medicine* **65**, 584-594.
265. Birch-Machin MA., Trunbull DM. (2001) Assaying mitochondrial respiratory complex activity in mitochondria isolated from human cells and tissues. *Methods in Cell Biology* **65**, 97-117.
266. Ritchie TK., Grinkova YV., Bayburt TH., Denisov IG., Zolnerciks JK., Atkins WM., Sligar SG., (2009) Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods in Enzymology* **464**, 211-231.
267. Atamna H., Machey J., Dhahbi JM. (2012) Mitochondrial pharmacology: electron transport chain bypass as strategies to treat mitochondrial dysfunction. *BioFactors* **38**, 158-166.
268. James LP., (2013) Chapter 20 – Acetaminophen: Pathology and Clinical Presentation of Hepatotoxicity, in Kaplowitz N., Deleve L., (Ed.) Drug-induced liver disease, 3rd ed. Amsterdam: Elsevier 331-341.